(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 June 2003 (12.06.2003)

PCT

(10) International Publication Number
WO 03/048320 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US02/38360

(22) International Filing Date:
27 November 2002 (27.11.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10/006,780 30 November 2001 (30.11.2001) US
10/086,935 28 February 2002 (28.02.2002) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier application:
US 10/086,935 (CIP)
Filed on 28 February 2002 (28.02.2002)

(71) Applicant *(for all designated States except US)*: **CYTOKINETICS, INC.** [US/US]; 280 East Grand Avenue, South San Francisco, CA 94080 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SAKOWICZ, Roman** [PL/US]; 1119 Nimitz Lane, Foster City, CA 94404 (US). **BERAUD, Christophe** [FR/US]; 761 Tehama Street, San Francisco, CA 94103 (US). **GUO, Jun** [CN/US]; 40326 San Sebastian Place, Fremont, CA 94539 (US). **FREEDMAN, Richard** [US/US]; 1012 Shoreline Drive, San Mateo, CA 94404 (US).

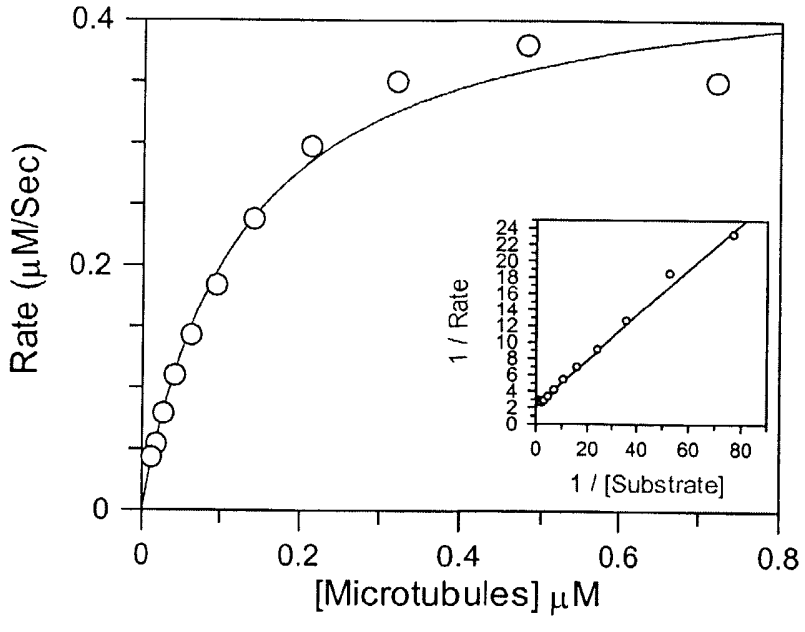
(74) Agents: **AUSENHUS, Scott, L.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111-3834 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND METHODS FOR ITS USE



(57) Abstract: The invention provides isolated nucleic acid and amino acid sequences of PfKinI-1, antibodies to PfKinI-1, methods of screening for PfKinI-1 modulators using biologically active PfKinI-1, and kits for screening for PfKinI-1 modulators.

WO 03/048320 A2

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND METHODS FOR ITS USE

Cross-Reference to Related Patent Applications

[0001] This application is a continuation-in part of U.S. Patent Application No. 10/086,935, filed February 28, 2002, which is a continuation-in-part of U.S. Patent Application number 10/006,780, filed November 30, 2001, both of which are incorporated herein by reference in their entirety for all purposes.

Field of the Invention

[0002] The invention provides isolated nucleic acid and amino acid sequences of *Plasmodium falciparum* KinI-1, methods of its detection, and screening for modulators using biologically active *Plasmodium falciparum* KinI-1, and kits therefor.

Background of the Invention

[0003] The kinesin superfamily is an extended family of related microtubule motor proteins. It can be classified into at least eight subfamilies based on primary amino acid sequence, domain structure, velocity of movement, and cellular function. This family is exemplified by "true" kinesin, which was first isolated from the axoplasm of squid, where it is believed to play a role in anterograde axonal transport of vesicles and organelles (*see, e.g.*, Goldstein, *Annu. Rev. Genet.* 27:319-351 (1993)). Kinesin uses ATP to generate force and directional movement associated with microtubules.

[0004] Mitotic kinesins are enzymes essential for assembly and function of the mitotic spindle, but are not generally part of other microtubule structures. Mitotic kinesins play essential roles during all phases of mitosis. These enzymes are "molecular motors" that translate energy released by hydrolysis of ATP into mechanical force which drives the directional movement of cellular cargoes along microtubules. The catalytic domain sufficient for this task is a compact structure of approximately 340 amino acids.

[0005] Within this functional group of kinesins resides a group of kinesins from several organisms that share significant sequence homology, the Kin I subfamily, that function to destabilize microtubule ends. These include *H. sapiens* MCAK (also known as mitotic centromere-associated kinesin or HsMCAK), *X. laevis* MCAK, and *C. griseus* MCAK. During

anaphase A, disjoined sister chromatids migrate poleward. This poleward movement is driven by kinetochores and is accompanied by the depolymerization of microtubules attached to the migrating chromatids. The kinesin MCAK plays an important role in this motility and may promote disassembly of microtubules attached to kinetochores of mitotic chromosomes.

[0006] The HsMCAK gene has a predicted 723 amino acid open reading frame, encoding a 81 kDa protein that shares 79.2% homology with hamster MCAK. Genes for the Xenopus and hamster homologs of MCAK have also been cloned and characterized. Defects in the function of these proteins would be expected to result in cell cycle arrest in mitosis.

[0007] Malaria is an acute or chronic disease caused by the presence of sporozoan parasites of the genus *Plasmodium* in the host's red blood cells (erythrocytes). Four different species can cause the disease in humans: *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale* and *Plasmodium malariae*. The major cause of malaria in humans is *P. falciparum* which infects 200 million to 400 million people every year, killing 1 to 4 million. Approximately 25 percent of all deaths of children in rural Africa between the ages of one and four years are caused by malaria.

[0008] The blood stage infection which is entirely responsible for the symptoms of malaria, begins with the entry of a merozoite into the erythrocyte. The intra-erythrocytic parasite develops through morphologically distinct ring (0-24 hours) and trophozoite (24-36 hours) stage to schizogony (36-48 hours), where mitosis occurs and 10-16 daughter merozoites are assembled. At the end of the schizogony, the infected erythrocyte ruptures and the released merozoites reinvade red cells to maintain the asexual cycle.

[0009] Merozites have very few microtubules. See, Bannister and Mitchell (1995) *Annals of Tropical Medicine and Parasitology* 89:105-111. These microtubules are arranged in a form termed the fMAST. See, Fowler *et al.* (1998) *Parasitology Today* 14:41. Disorganization of these microtubules has been shown to inhibit invasion of the merozites into the red cells. See, Bejon *et al.* (1996) *Parasitology* 114:1-6. On the other hand, stabilization of the microtubules, for example with a drug such as Taxol™, does not interfere with invasion. See, Fowler et al. (1998) *Parasitology* 117:425-433. It has been suggested that fMAST may in fact act as a track along which kinesin motors run in the invasion process.

[0010] The discovery of a new kinesin motor protein which is the *P. falciparum* ortholog of human MCAK, is a microtubule-stimulated ATPase and depolymerizes microtubules. The polynucleotides encoding it satisfy a need in the art by providing new compositions that are useful in the diagnosis, prevention, and treatment of malaria.

Summary of the Invention

[0011] The present invention is based on the discovery of a new kinesin motor protein, *P. falciparum* KinI-1, the polynucleotide encoding *P. falciparum* KinI-1, and the use of these compositions for the diagnosis, treatment, or prevention of malaria.

[0012] In one aspect, the invention provides an isolated nucleic acid sequence encoding a kinesin superfamily motor protein, wherein the motor protein has the following properties: (i) the protein has microtubule stimulated ATPase activity and/or depolymerizes microtubules; and (ii) the protein comprises a sequence that has greater than 90% amino acid sequence identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 as measured using a sequence comparison algorithm. In one embodiment, the protein further specifically binds to polyclonal antibodies raised against SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

[0013] In one embodiment, the nucleic acid encodes PfKinI-1 or a fragment thereof. In another embodiment, the nucleic acid encodes SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12. In another embodiment, the nucleic acid comprises a nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11.

[0014] In one aspect, the nucleic acid comprises a sequence which encodes an amino acid sequence which has greater than 70% sequence identity with SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12, preferably greater than 80%, more preferably greater than 90%, more preferably greater than 95% or, in another embodiment, has 98 or 100% sequence identity with SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

[0015] In one embodiment, the nucleic acid comprises a sequence which has greater than 55 or 60% sequence identity with SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11, preferably greater than 70%, more preferably greater than 80%, more preferably greater than 90 or 95% or, in another embodiment, has 98 or 100% sequence identity with SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11. In another embodiment provided herein, the nucleic acid hybridizes under stringent conditions to a nucleic acid having a sequence or complementary sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11.

[0016] In another aspect, the invention provides an expression vector comprising a nucleic acid encoding a kinesin superfamily motor protein, wherein the motor protein has the following properties: (i) the protein has microtubule stimulated ATPase activity; and (ii) the

protein comprises a sequence that has greater than 70% amino acid sequence identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 as measured using a sequence comparison algorithm. The invention further provides a host cell transfected with the vector.

[0017] In another aspect, the invention provides an isolated kinesin superfamily motor protein, wherein the protein has one or more of the properties described above. In one embodiment, the protein specifically binds to polyclonal antibodies generated against a motor domain, tail domain or other fragment of PfKinI-1. In another embodiment, the protein comprises an amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

[0018] In one aspect, the protein provided herein comprises an amino acid sequence which has greater than 70% sequence identity with SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12, preferably greater than 80%, more preferably greater than 90%, more preferably greater than 95% or, in another embodiment, has 98 or 100% sequence identity with SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

[0019] The invention features a substantially purified polypeptide comprising the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 or a fragment thereof and more particularly, the motor domain of the amino acid sequence of SEQ ID NO:2 or a fragment thereof, optionally coupled to one or more epitope tags.

[0020] In one embodiment, the present invention provides a method of identifying a candidate agent as a modulator of the activity of a target protein. The method comprises adding a candidate agent to a mixture comprising a target protein which directly or indirectly produces ADP or phosphate, under conditions that normally allow the production of ADP or phosphate. The method further comprises subjecting the mixture to a reaction that uses said ADP or phosphate as a substrate under conditions that normally allow the ADP or phosphate to be utilized and determining the level of activity of the reaction as a measure of the concentration of ADP or phosphate. A change in the level between the presence and absence of the candidate agent indicates a modulator of the target protein.

[0021] The phrase "use ADP or phosphate" means that the ADP or phosphate are directly acted upon by detection reagents. In one case, the ADP, for example, can be hydrolyzed or can be phosphorylated. As another example, the phosphate can be added to

another compound. As used herein, in each of these cases, ADP or phosphate is acting as a substrate.

[0022] Preferably, the target protein either directly or indirectly produces ADP or phosphate and comprises a motor domain. More preferably, the target protein comprises a kinesin superfamily motor protein as described above and most preferably, the target protein comprises PfKinI-1 or a fragment thereof.

[0023] Also provided are modulators of the target protein including agents for the treatment of malaria. The agents and compositions provided herein can be used in variety of applications which include the formulation of sprays, powders, and other compositions. Also provided herein are methods of treating malaria, for treating disorders associated with PfKinI-1 activity, and for inhibiting PfKinI-1.

[0024] Antibodies that specifically bind to a protein having a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10; and SEQ ID NO:12 are also provided. The antibodies can be polyclonal, monoclonal, or recombinant antibodies, or fragments thereof, for instance.

[0025] Pharmaceutical compositions useful for treating diseases associated with *P. falciparum* infection are also provided. These compositions typically include an agent that inhibits the microtubule-stimulated ATPase and/or microtubule depolymerization activity of a protein having a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10; and SEQ ID NO:12 and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026] FIG. 1A shows plots of the kinetics of ATP hydrolysis by PfKinI-1 in the presence of varying concentrations of microtubules.

[0027] FIG. 1B is a gel depicting microtubule depolymerization by PfKinI-1. In this experiment, Taxol™-stabilized microtubules were incubated in the presence of ATP and substoichiometric concentrations of PfKinI-1. Tubulin released from microtubule polymers was separated by ultracentrifugation. Supernatant (S) and pellet (P) fractions were analyzed by SDS-PAGE. Presence of tubulin in the supernatant is indicative of tubulin depolymerization.

[0028] FIG. 2A is a plot showing PfKinI-1-induced microtubule depolymerization. Depolymerization was measured by following the decrease in DAPI fluorescence as tubulin dimers dissociate from preformed protofilaments.

[0029] FIG. 2B is a chart showing the slope corresponding to the rate of depolymerization reactions catalyzed by PfKinI-1 relative to control reactions lacking enzyme or ATP.

BRIEF DESCRIPTION OF THE SEQUENCE LISTING

[0030] The Sequence Listing, which is incorporated herein by reference in its entirety, provides exemplary sequences including polynucleotide sequences, SEQ ID NOs: 1, 3, 5, 7, 9, 11, and 13, and polypeptide sequences, SEQ ID NOs: 2, 4, 6, 8, 10, 12 or 14. Each sequence is identified by a sequence identification number (SEQ ID NO).

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0031] "**ADP**" refers to adenosine diphosphate and also includes ADP analogs, including, but not limited to, deoxyadenosine diphosphate (dADP) and adenosine analogs.

[0032] "**Antibody**" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively. The term antibody also includes antibody fragments either produced by the modification of whole antibodies or those synthesized de novo using recombinant DNA methodologies.

[0033] An "**anti-PfKinI-1**" antibody is an antibody or antibody fragment that specifically binds a polypeptide encoded by the PfKinI-1 gene, cDNA, or a subsequence thereof.

[0034] "**Biologically active**" target protein refers to a target protein that has one or more of kinesin protein's biological activities, including, but not limited to microtubule stimulated ATPase activity, as tested, e.g., in an ATPase assay. Biological activity can also be demonstrated in a microtubule gliding assay, a microtubule binding assay, or a microtubule depolymerization assay. "ATPase activity" refers to ability to hydrolyze ATP. Other activities include polymerization/depolymerization (effects on microtubule dynamics), binding to other proteins of the spindle, binding to proteins involved in cell-cycle control, or serving as a

substrate to other enzymes, such as kinases or proteases and specific kinesin cellular activities, such as chromosome congregation, axonal transport, etc.

[0035] **"Biological sample"** as used herein is a sample of biological tissue or fluid that contains a target protein or a fragment thereof or nucleic acid encoding a target protein or a fragment thereof. Biological samples may also include sections of tissues such as frozen sections taken for histological purposes. A biological sample comprises at least one cell, preferably plant or vertebrate. Embodiments include cells obtained from a eukaryotic organism, preferably eukaryotes such as fungi, plants, insects, protozoa, birds, fish, reptiles, and preferably a mammal such as rat, mice, cow, dog, guinea pig, or rabbit, and most preferably a primate such as chimpanzees or humans.

[0036] A **"comparison window'** includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 25 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, e.g., by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.* 2:482 (1981), by the global alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.* 48:443 (1970), by the search for similarity methods of Pearson & Lipman, *Proc. Natl. Acad. Sci. USA* 85:2444 (1988) and Altschul et al. *Nucleic Acids Res.* 25(17): 3389-3402 (1997), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and BLAST in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (see, e.g., Ausubel et al., supra).

[0037] This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always ,)). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity

X from its maximum achieved value; the cumulative score goes to zero, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. For identifying whether a nucleic acid or polypeptide is within the scope of the invention, the default parameters of the BLAST programs are suitable. The BLASTN program (for nucleotide sequences) uses as defaults a word length (W) of 11, an expectation (E) of 10, M =5, N = -4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a word length (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix. The TBLATN program (using protein sequence for nucleotide sequence) uses as defaults a word length (W) of 3, an expectation (E) of 10, and a BLOSUM 62 scoring matrix. See, Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915.

[0038] In addition to calculating percent sequence identity, the BLAST algorithm also performs a statistical analysis of the similarity between two sequences (see, e.g., Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5787). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.1, more preferably less than about 0.01, and most preferably less than about 0.001.

[0039] Another example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments. It can also plot a dendrogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.* 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, CABIOS 5:151-153 (1989). As a general rule, PILEUP can align up to 500 sequences, with any single sequence in the final alignment restricted to a maximum length of 7,000 characters.

[0040] The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster can then be aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences can be aligned by a simple extension of the pairwise alignment of two individual sequences. A series of such pairwise alignments that includes increasingly dissimilar sequences and clusters of sequences at each iteration produces the final alignment.

[0041] **"Variant"** applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, conservatively modified variants refers to those nucleic acids which encode identical or essentially identical amino acid sequences, or where the nucleic acid does not encode an amino acid sequence, to essentially identical sequences. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given protein. For instance, the codons GCA, GCC, GCG and GCT all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are "silent variations," which are one species of conservatively modified variations. Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each degenerate codon in a nucleic acid can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

[0042] Also included within the definition of target proteins of the present invention are amino acid sequence variants of wild-type target proteins. These variants fall into one or more of three classes: substitutional, insertional or deletional variants. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the target protein, using cassette or PCR mutagenesis or other techniques well known in the art, to produce DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. Variant target protein fragments having up to about 100-150 amino acid residues may be prepared by in vitro synthesis using established techniques. Amino acid sequence variants are characterized by the predetermined nature of the variation, a feature that sets them apart from naturally occurring allelic or interspecies variation of the target protein amino acid sequence. The variants typically exhibit the same qualitative biological activity as the naturally occurring analogue, although variants can also be selected which have modified characteristics.

[0043] Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to about 20 amino acids, although considerably longer insertions may be tolerated. Deletions range from about 1 to about 20 residues, although in some cases, deletions may be much longer.

[0044] Substitutions, deletions, and insertions or any combinations thereof may be used to arrive at a final derivative. Generally, these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger characteristics may be tolerated in certain circumstances.

[0045] The following six groups each contain amino acids that are conservative substitutions for one another:

Alanine (A), Serine (S), Threonine (T);

Aspartic acid (D), Glutamic acid (E);

Asparagine (N), Glutamine (Q);

Arginine (R), Lysine (K);

Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and

Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

(*see, e.g.*, Creighton, *Proteins* (1984)).

[0046] "**Cytoskeletal component**" denotes any molecule that is found in association with the cellular cytoskeleton, that plays a role in maintaining or regulating the structural integrity of the cytoskeleton, or that mediates or regulates motile events mediated by the cytoskeleton. Cytoskeletal component is intended to include cytoskeletal polymers (e.g., actin filaments, microtubules, intermediate filaments, myosin fragments), molecular motors (e.g., kinesins, myosins, dyneins), cytoskeleton associated regulatory proteins (e.g., tropomysin, alpha-actinin) and cytoskeletal associated binding proteins (e.g., microtubules associated proteins, actin binding proteins).

[0047] "**Cytoskeletal function**" refers to biological roles of the cytoskeleton, including but not limited to the providing of structural organization (e.g., microvilli, mitotic spindle) and the mediation of motile events within the cell (e.g., muscle contraction, mitotic chromosome movements, contractile ring formation and function, pseudopodal movement, active cell surface deformations, vesicle formation and translocation.)

[0048] A "**diagnostic**" as used herein is a compound, method, system, or device that assists in the identification and characterization of a health or disease state. The diagnostic can be used in standard assays as is known in the art.

[0049] An "**expression vector**" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements that permit transcription of a particular nucleic acid in a host cell. The expression vector can be part of a plasmid, virus, or nucleic acid fragment. Typically, the expression vector includes a nucleic acid to be transcribed operably linked to a promoter.

[0050] "**High throughput screening**" as used herein refers to an assay which provides for multiple candidate agents or samples to be screened simultaneously. As further described below, examples of such assays may include the use of microtiter plates which are especially

convenient because a large number of assays can be carried out simultaneously, using small amounts of reagents and samples.

[0051] By "**host cell**" is meant a cell that contains an expression vector and supports the replication or expression of the expression vector. Host cells may be prokaryotic cells such as *E. coli*, or eukaryotic cells such as yeast, insect, amphibian, or mammalian cells such as CHO, HeLa and the like, or plant cells. Both primary cells and cultured cell lines are included in this definition.

[0052] The phrase "**hybridizing specifically to**" or "**specific hybridization**" or "**selectively hybridize to**", refers to the binding, duplexing, or hybridizing of a nucleic acid molecule preferentially to a particular nucleotide sequence under stringent conditions when that sequence is present in a complex mixture (e.g., total cellular) DNA or RNA.

[0053] The term "**stringent conditions**" refers to conditions under which a probe will hybridize preferentially to its target subsequence, and to a lesser extent to, or not at all to, other sequences. Astringent hybridization wash conditions in the context of nucleic acid hybridization experiments such as Southern and Northern hybridizations are sequence dependent, and are different under different environmental parameters. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) Laboratory Techniques in Biochemistry and Molecular Biology – Hybridization with Nucleic Acid Probes, Part I, Chapter 2: An overview of principles of hybridization and the strategy of nucleic acid probe assays, Elsevier, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.05 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides) and at least about 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Very stringent conditions are selected to be equal to the T_m for a particular probe.

[0054] An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or Northern blot is 50% formamide with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is

0.15 NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (see, Sambrook et al. (1989) Molecular Cloning – A Laboratory Manual (2nd Ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY. Often a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example of a medium stringency wash for a duplex of, e.g., more than 100 nucleotides, is 1x SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, e.g., more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids which do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, e.g., when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

[0055] The terms **"identical"** or percent **"identity"**, in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Preferably, the percent identity exists over a region of the sequence that is at least about 25 amino acids in length, more preferably over a region that is 50 or 100 amino acids in length. This definition also refers to the complement of a test sequence, provided that the test sequence has a designated or substantial identity to a reference sequence. Preferably, the percent identity exists over a region of the sequence that is at least about 25 nucleotides in length, more preferably over a region that is 50 or 100 nucleotides in length.

[0056] When percentage of sequence identity is used in reference to proteins or peptides, it is recognized that residue positions that are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g.,. charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Means for making this adjustment are well known to those of skill in the art. The scoring of conservative substitutions can be calculated according to, e.g., the algorithm of Meyers & Millers, *Computer Applic. Biol. Sci.* 4:11-17 (1988), e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, California).

[0057] The terms "**isolated**", "**purified**", or "**biologically pure**" refer to material that is substantially or essentially free from components which normally accompany it as found in its native state. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified. In an isolated gene, the nucleic acid of interest is separated from open reading frames which flank the gene of interest and encode proteins other than the protein of interest. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure.

[0058] A "**label**" is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include fluorescent proteins such as green, yellow, red or blue fluorescent proteins, radioisotopes such as ^{32}P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available (e.g., the polypeptide of SEQ ID NO:2 can be made detectable, e.g., by incorporating a radio-label into the peptide, and used to detect antibodies specifically reactive with the peptide).

[0059] A "**labeled nucleic acid probe or oligonucleotide**" is one that is bound, either covalently, through a linker, or through ionic, van der Waals, or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

[0060] "**Moderately stringent conditions**" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 μg/mL denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

[0061] **"Modulators," "inhibitors,"** and **"activators of a target protein"** refer to modulatory molecules identified using *in vitro* and *in vivo* assays for target protein activity. Such assays include ATPase activity, microtubule gliding, microtubule depolymerizing activity, and binding activity such as microtubule binding activity or binding of nucleotide analogs. Samples or assays are treated with a candidate agent at a test and control concentration. The control concentration can be zero. If there is a change in target protein activity between the two concentrations, this change indicates the identification of a modulator. A change in activity, which can be an increase or decrease, is preferably a change of at least 20% to 50%, more preferably by at least 50% to 75%, more preferably at least 75% to 100%, and more preferably 150% to 200%, and most preferably is a change of at least 2 to 10 fold compared to a control. Additionally, a change can be indicated by a change in binding specificity or substrate.

[0062] **"Molecular motor"** refers to a molecule that utilizes chemical energy to generate mechanical force. According to one embodiment, the molecular motor drives the motile properties of the cytoskeleton.

[0063] The phrase **"motor domain"** refers to the domain of a target protein that confers membership in the kinesin superfamily of motor proteins through a sequence identity of approximately 35-45% identity to the motor domain of true kinesin.

[0064] The term **"nucleic acid"** refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides which have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences as well as the sequence explicitly indicated. For example, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer et al., *Nucleic Acid Res.* 19:5081 (1991); Ohtsuka et al., *J. Biol. Chem.* 260)2605-2608 (1985); Cassol et al. 1992; Rossolini et al. *Mol. Cell. Probes* 8:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

[0065] **"Nucleic acid probe or oligonucleotide"** is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (i.e., A, G, C, or T) or modified bases.

In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidine complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

[0066] The terms **"polypeptide"**, **"peptide"** and **"protein"** are used interchangeably herein to refer to a polymer of amino acid residues. The terms apply to amino acid polymers in which one or more amino acid residues is an artificial chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers. A target protein comprises a polypeptide demonstrated to have at least microtubule stimulated ATPase activity. Amino acids may be referred to herein by either their commonly known three letter symbols or by Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes, i.e., the one-letter symbols recommended by the IUPAC-IUB.

[0067] A **"promoter"** is defined as an array of nucleic acid control sequences that direct transcription of a nucleic acid. As used herein, a promoter includes necessary nucleic acid sequences near the start site of transcription, such as, in the case of a polymerase II type promoter, a TATA box element. A promoter also optionally includes distal enhancer or repressor elements which can be located as much as several thousand base pairs from the start site of transcription. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is under environmental or developmental regulation. The term "operably linked" refers to a functional linkage between a nucleic acid expression control sequence (such as a promoter, or array of transcription factor binding sites) and a second nucleic acid sequence, wherein the expression control sequence directs transcription of the nucleic acid corresponding to the second sequence.

[0068] The phrase **"specifically (or selectively) binds"** to an antibody or **"specifically (or selectively) immunoreactive with,"** when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions,

the specified antibodies bind to a particular protein at least two times the background and do not substantially bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to PfKinI-1 with the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 can be selected to obtain only those antibodies that are specifically immunoreactive with PfKinI-1 and not with other proteins, except for polymorphic variants, orthologs, alleles, and closely related homologues of PfKinI-1. This selection may be achieved by subtracting out antibodies that cross react with molecules, for example, such as *C. elegans* unc-104 and human Kif1A. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual* (1988), for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity). Typically a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

[0069] The phrase "**selectively associates with**" refers to the ability of a nucleic acid to "selectively hybridize" with another as defined above, or the ability of an antibody to "selectively (or specifically) bind to a protein, as defined above.

[0070] "**Test composition**" (used interchangeably herein with "**candidate agent**" and "**test compound**" and "**test agent**") refers to a molecule or composition whose effect on the interaction between one or more cytoskeletal components it is desired to assay. The "test composition" can be any molecule or mixture of molecules, optionally in a carrier.

[0071] A "**therapeutic**" as used herein refers to a compound which is believed to be capable of modulating the cytoskeletal system *in vivo* which can have application in both human and animal disease.

[0072] A "**baseline value**" or "**standard value**" refers to a value (or range of values) against which an experimental or determined value (e.g., one determined for a patient sample as part of a diagnostic test) is compared in an assay or diagnostic test. Thus, in the case of PfKinI-1 activity or expression level, for example, the baseline value can be a value for PfKinI-1 activity or expression for a sample obtained from the same individual at different time points. In other instances, the baseline value is a value determined for a control reaction, cell or individual, or a statistical value (e.g., an average or mean) established for a population of control cells or individuals. In certain tests for PfKinI-1 activity or expression, the control can

be a cell, individual or populations thereof which is not infected with *P. falciparum*, for example. The population that serves as a control can vary in size, having as few as a single member, but potentially including tens, hundreds, thousands, tens of thousands or more individuals. When the control is a large population, the baseline value can be a statistical value determined from individual values for each member or a value determined from the control population as an aggregate (e.g., a value measured for a population of cells within a well).

[0073] A difference is typically considered to be "**statistically significant**" in an assay or diagnostic test if the probability of the observed difference occurring by chance (the p-value) is less than some predetermined level. As used herein a "statistically significant difference" refers to a p-value that is < 0.05, preferably < 0.01 and most preferably < 0.001.

Nucleic Acids and The Target Protein

[0074] The present invention provides for the first time a nucleic acid encoding PfKinI-1. This protein is a member of the kinesin superfamily of motor proteins. More specifically, PfKinI-1 shares approximately 50% identity to a fragment of HsKinI-3 (also known as AL3636622). See, PCT Publication WO 02/26929, which is incorporated herein by reference.

[0075] In one aspect, PfKinI-1 can be defined by having at least one or preferably more than one of the following functional and structural characteristics. Functionally, PfKinI-1 has microtubule-stimulated ATPase activity. PfKinI-1 activity can also be described in terms of its ability to depolymerize microtubules.

[0076] The novel nucleotides sequences provided herein encode PfKinI-1 or fragments thereof. Thus, in one aspect, the nucleic acids provided herein are defined by the novel proteins provided herein. The protein provided herein comprises an amino acid sequence which has one or more of the following characteristics: greater than 70% sequence identity with SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12, preferably greater than 80% or 85%, more preferably greater than 90%, 92% or 94%, more preferably greater than 95% or, in another embodiment, has 98 or 100% sequence identity with SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12. The region of percent identity extends over a region of at least 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, or 300 contiguous amino acids. In other instances, the region extends over the full length of the amino acid sequence set forth in SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12. As described above, when describing the nucleotide in terms of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID

NO:9, or SEQ ID NO:11, the sequence identity may be slightly lower due to the degeneracy in the genetic code.

[0077] The invention also includes fragments of the nucleotide sequence having at least 10, 15, 20, 25, 30, 40, 50, 100, 200, 300, 400, 500, 1000 or 2000 contiguous nucleotides from SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11, or a degenerate form thereof. Some fragments include the motor domain or variants thereof, e.g., SEQ ID NOs:4, 6, 8, or 12, which occurs approximately between positions 4 or 5 to about positions 334 or 335 of the amino acid SEQ ID NO:2 (determined by sequence comparison of the motor domain of other kinesins). Some such fragments can be used as hybridization probes or primers. Unless otherwise apparent from the context, reference to nucleotide sequences shown in the sequence listing can refer to the sequence shown, its perfect complement or a duplex of the two strands. Also included within the definition of the target proteins are amino acid sequence variants of wild-type target proteins.

[0078] Portions of the PfKinI-1 nucleotide sequence may be used to identify polymorphic variants, orthologs, alleles, and homologues of PfKinI-1. This identification can be made *in vitro*, e.g., under stringent hybridization conditions and sequencing, or by using the sequence information in a computer system for comparison with other nucleotide sequences. Sequence comparison can be performed using any of the sequence comparison algorithms discussed below, with PILEUP as a preferred algorithm.

[0079] As will be appreciated by those in the art, the target proteins can be made in a variety of ways, including both synthesis de novo and by expressing a nucleic acid encoding the protein.

[0080] Target proteins of the present invention may also be modified in a way to form chimeric molecules comprising a fusion of a target protein with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino or carboxyl terminus of the target protein. Provision of the epitope tag enables the target protein to be readily detected, as well as readily purified by affinity purification. Various tag epitopes are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 (see, Field et al. (1988) *Mol. Cell. Biol.* 8:2159); the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto (see, Evans et al., (1985) *Molecular and Cellular Biology*, 5:3610); and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (see, Paborsky et al., (1990) *Protein Engineering*, 3:547). Other tag polypeptides include the Flag-peptide (see, Hopp et al. (1988) *BioTechnology* 6:1204); the KT3 epitope

peptide (see, Martine et al. (1992) *Science*, 255:192); tubulin epitope peptide (see, Skinner (1991) *J. Biol. Chem.* 266:15173); and the T7 gene 10 protein peptide tag (see, Lutz-Freyermuth et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6393.

[0081] The biological activity of any of the peptides provided herein can be routinely confirmed by the assays provided herein such as those which assay ATPase activity or microtubule binding activity. In one embodiment, polymorphic variants, alleles, and orthologs, homologues of PfKinI-1 are confirmed by using a ATPase or microtubule binding assays as known in the art.

[0082] The isolation of biologically active PfKinI-1 for the first time provides a means for assaying for modulators of this kinesin superfamily protein. Biologically active PfKinI-1 is useful for identifying modulators of PfKinI-1 or fragments thereof and kinesin superfamily members using *in vitro* assays such as microtubule gliding assays, ATPase assays (Kodama *et al., J. Biochem.* 99:1465-1472 (1986); Stewart *et al., Proc. Nat'l Acad. Sci. USA* 90:5209-5213 (1993)), and binding assays including microtubule binding assays (Vale *et al.*, *Cell* 42:39-50 (1985)). *In vivo* assays and uses are provided herein as well. Also provided herein are methods of identifying candidate agents which bind to PfKinI-1 and portions thereof.

[0083] Some portions or fragments of PfKinI-1 include at least 7, 10, 15, 20, 35, 50, 100, 250, 300, 350, 500, or 1000 contiguous amino acids from the sequences listed in SEQ ID NOS:2, 4, 6, 8, 10 or 12 Some fragments contain fewer than 1000, 500, 250, 100, or 50 contiguous amino acids from the sequences shown in SEQ ID NOS:2, 4, 6, 8, 10 or 12. For example, exemplary fragments include fragments having 15-50 amino acids or 100-500 amino acids. Some fragments include a motor domain. The motor domain runs from about amino acid 4 to about amino acid 401 of SEQ ID NO:2. Such fragments typically include this span, or an active portion thereof. Some fragments include a ligand binding domain of PfKinI-1. Nucleic acids encoding such fragments are also included in the invention.

[0084] As further described herein, a wide variety of assays, therapeutic and diagnostic methods are provided herein which utilize the novel compounds described herein. The uses and methods provided herein, as further described below have *in vivo*, *in situ*, and *in vitro* applications, and can be used in medicinal, veterinary, agricultural and research based applications.

Isolation of the gene encoding PfKinI-1

General Recombinant DNA Methods

[0085] This invention relies on routine techniques in the field of recombinant genetics. Basic texts disclosing the general methods of use in this invention include Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2nd ed. 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); and *Current Protocols in Molecular Biology* (Ausubel *et al.*, eds., 1994)).

[0086] For nucleic acids, sizes are given in either kilobases (kb) or base pairs (bp). These are estimates derived from agarose or acrylamide gel electrophoresis, from sequenced nucleic acids, or from published DNA sequences. For proteins, sizes are given in kilodaltons (kDa) or amino acid residue numbers. Proteins sizes are estimated from gel electrophoresis, from mass spectroscopy, from sequenced proteins, from derived amino acid sequences, or from published protein sequences.

[0087] Oligonucleotides that are not commercially available can be chemically synthesized according to the solid phase phosphoramidite triester method first described by Beaucage & Caruthers, *Tetrahedron Letts.* 22:1859-1862 (1981), using an automated synthesizer, as described in Van Devanter *et al., Nucleic Acids Res.* 12:6159-6168 (1984). Purification of oligonucleotides is by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson & Reanier, *J. Chrom.* 225:137-149 (1983).

[0088] The sequence of the cloned genes and synthetic oligonucleotides can be verified after cloning using, e.g., the chain termination method for sequencing double-stranded templates of Wallace *et al., Gene* 16:21-26 (1981).

Cloning methods for the isolation of nucleotide sequences encoding PfKinI-1

[0089] In general, the nucleic acid sequences encoding PfKinI-1 and related nucleic acid sequence homologs are cloned from cDNA and genomic DNA libraries or isolated using amplification techniques with oligonucleotide primers. Alternatively, expression libraries can be used to clone PfKinI-1 and PfKinI-1 homologues by detected expressed homologues immunologically with antisera or purified antibodies made against PfKinI-1 that also recognize and selectively bind to the PfKinI-1 homologue. Finally, amplification techniques using primers can be used to amplify and isolate PfKinI-1 from DNA or RNA. Amplification techniques using degenerate primers can also be used to amplify and isolate PfKinI-1 homologues. Amplification techniques using primers can also be used to isolate a nucleic acid encoding PfKinI-1. These primers can be used, e.g., to amplify a probe of several hundred nucleotides, which is then used to screen a library for full-length PfKinI-1.

[0090] Appropriate primers and probes for identifying the gene encoding homologues of PfKinI-1 in other species are generated from comparisons of the sequences provided herein. As described above, antibodies can be used to identify PfKinI-1 homologues. For example, antibodies made to the motor domain of PfKinI-1 or to the whole protein are useful for identifying PfKinI-1 homologues.

[0091] To make a cDNA library, one should choose a source that is rich in the mRNA of choice, e.g., PfKinI-1. Expression is dramatically increased in testes and fetal liver. The mRNA is then made into cDNA using reverse transcriptase, ligated into a recombinant vector, and introduced into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known (see, e.g., Gubler & Hoffman, *Gene* 25: 263-269); Sambrook et al., supra; Ausubel et al., supra).

[0092] For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of about 12-20 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro*. Recombinant phage are analyzed by plaque hybridization as described in Benton & Davis, Science 196:180-182 (1977). Colony hybridization is read out as generally described in Grunstein et al., *Proc. Natl. Acad. Sci. USA*, 72:3961-3965 (1975).

[0093] An alternative method of isolating PfKinI-1 nucleic acid and its homologues combines the use of synthetic oligonucleotide primers and amplification of an RNA or DNA template (see U.S. Patents 4,683,195 and 4,683,202; PCR Protocols: A Guide to Methods and Applications (Innis et al., eds. 1990)). Methods such as polymerase chain reaction and ligase chain reaction can be used to amplify nucleic acid sequences of PfKinI-1 directly from mRNA, from cDNA, from genomic libraries or cDNA libraries. Degenerate oligonucleotides can be designed to amplify PfKinI-1 homologues using the sequences provided herein. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of PfKinI-1 encoding mRNA in physiological samples, for nucleic sequencing or for other purposes. Genes amplified by the PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

[0094] Gene expression of PfKinI-1 can also be analyzed by techniques known in the art, e.g., reverse transcription and amplification of mRNA, isolation of total RNA or poly A +

RNA, blotting, dot blotting, in situ hybridization, RNase protection, quantitative PCR, and the like.

[0095] Synthetic oligonucleotides can be used to construct recombinant PfKinI-1 genes for use as probes or for expression of protein. This method is performed using a series of overlapping oligonucleotides usually 40-120 bp in length, representing both the sense and antisense strands of the gene. These DNA fragments are then annealed, ligated and cloned. Alternatively, amplification techniques can be used with precise primers to amplify a specific subsequence of the PfKinI-1 gene. The specific subsequence is then ligated into an expression vector.

[0096] The gene for PfKinI-1 is typically cloned into intermediate vectors before transformation into prokaryotic or eukaryotic cells for replication and/or expression. The intermediate vectors are typically prokaryote vectors or shuttle vectors.

Expression in prokaryotes and eukaryotes

[0097] To obtain high level expression of a cloned gene, such as those cDNAs encoding PfKinI-1, it is important to construct an expression vector that contains a strong promoter to direct transcription, a transcription/translation terminator, and if for a nucleic acid encoding a protein, a ribosome binding site for translational initiation. Suitable bacterial promoters are well known in the art and described, e.g., in Sambrook *et al.* and Ausubel *et al.* Bacterial expression systems for expressing the PfKinI-1 protein are available in, e.g., *E. coli, Bacillus sp.,* and *Salmonella* (Palva *et al.*, *Gene* 22:229-235 (1983); Mosbach *et al.*, *Nature* 302:543-545 (1983). Kits for such expression systems are commercially available. Eukaryotic expression systems for mammalian cells, yeast, and insect cells are well known in the art and are also commercially available. The pET expression system (Novagen) is a preferred prokaryotic expression system.

[0098] The promoter used to direct expression of a heterologous nucleic acid depends on the particular application. The promoter is preferably positioned about the same distance from the heterologous transcription start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

[0099] In addition to the promoter, the expression vector typically contains a transcription unit or expression cassette that contains all the additional elements required for the expression of the PfKinI-1 encoding nucleic acid in host cells. A typical expression cassette thus contains a promoter operably linked to the nucleic acid sequence encoding PfKinI-1 and

signals required for efficient polyadenylation of the transcript, ribosome binding sites, and translation termination. The nucleic acid sequence encoding PfKinI-1 may typically be linked to a cleavable signal peptide sequence to promote secretion of the encoded protein by the transformed cell. Such signal peptides would include, among others, the signal peptides from tissue plasminogen activator, insulin, and neuron growth factor, and juvenile hormone esterase of *Heliothis virescens*. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

[00100] In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

[00101] The particular expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic or prokaryotic cells may be used. Standard bacterial expression vectors include plasmids such as pBR322 based plasmids, pSKF, pET23, and fusion expression systems such as GST and LacZ. Epitope tags can also be added to recombinant proteins to provide convenient methods of isolation, e.g., c-myc or histidine tags.

[00102] Expression vectors containing regulatory elements from eukaryotic viruses are typically used in eukaryotic expression vectors, e.g., SV40 vectors, cytomegalovirus vectors, papilloma virus vectors, and vectors derived from Epstein Bar virus. Other exemplary eukaryotic vectors include pMSG, pAV009/A^+, pMTO10/A^+, pMAMneo-5, baculovirus pDSVE, and any other vector allowing expression of proteins under the direction of the SV40 early promoter, SV40 late promoter, CMV promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

[00103] Some expression systems have markers that provide gene amplification such as thymidine kinase, hygromycin B phosphotransferase, and dihydrofolate reductase. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a PfKinI-1 encoding sequence under the direction of the polyhedrin promoter or other strong baculovirus promoters.

[00104] The elements that are typically included in expression vectors also include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of

the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are preferably chosen such that they do not interfere with the replication of the DNA in eukaryotic cells, if necessary.

[00105] Standard transfection or transformation methods are used to produce bacterial, mammalian, yeast or insect cell lines that express large quantities of PfKinI-1 protein, which are then purified using standard techniques (*see, e.g.*, Colley *et al., J. Biol. Chem.* 264:17619-17622 (1989); *Guide to Protein Purification*, in *Methods in Enzymology*, vol. 182 (Deutscher ed., 1990)).

[00106] Transformation of eukaryotic and prokaryotic cells are performed according to standard techniques (*see, e.g.*, Morrison, *J. Bact.*, 132:349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology*, 101:347-362 (Wu *et al.*, eds, 1983). Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*, Sambrook *et al., supra*). It is only necessary that the particular genetic engineering procedure used be capable of successfully introducing at least one gene into the host cell capable of expressing PfKinI-1.

[00107] After the expression vector is introduced into the cells, the transfected cells are cultured under conditions favoring expression of PfKinI-1, which is recovered from the culture using standard techniques identified below.

Purification of PfKinI-1 Protein

[00108] Either naturally occurring or recombinant PfKinI-1 can be purified for use in functional assays. In a preferred embodiment, the target proteins are purified for use in the assays to provide substantially pure samples. Alternatively, the target protein need not be substantially pure as long as the sample comprising the target protein is substantially free of other components that can contribute to the production of ADP or phosphate.

[00109] The target proteins may be isolated or purified in a variety of ways known to those skilled in the art depending on what other components are present in the sample. Standard purification methods include electrophoretic, molecular, immunological, and chromatographic techniques, including ion exchange, hydrophobic, affinity, and reverse-phase HPLC chromatography, chromatofocussing, selective precipitation with such substances as ammonium

sulfate; and others (see, e.g., Scopes, Protein Purification: Principles and Practice (1982); U.S. Patent No. 4,673,641; Ausubel et al. supra; and Sambrook et al., supra). For example, the target protein can be purified using a standard anti-target antibody column. Ultrafiltration and diafiltration techniques, in conjunction with protein concentration, are also useful. A preferred method of purification utilizes Ni-NTA agarose (Qiagen).

[00110] The expressed protein can be purified by standard chromatographic procedures to yield a purified, biochemically active protein. The activity of any of the peptides provided herein can be routinely confirmed by the assays provided herein such as those which assay ATPase activity or microtubule binding activity. Biologically active target protein is useful for identifying modulators of target protein or fragments thereof and kinesin superfamily members using *in vitro* assays such as microtubule gliding assays, ATPase assays (Kodama et al., *J. Biochem.* 99:1465-1472 (1986); Stewart et al., *Proc. Nat'l Acad. Sci. USA* 90:5209-5213 (1993)), and binding assays including microtubule binding assays (Vale et al., *Cell* 42:39-50 (1985)), as described in detail below.

Purification of PfKinI-1 from recombinant bacteria

[00111] Recombinant proteins are expressed by transformed bacteria in large amounts, typically after promoter induction; but expression can be constitutive. Promoter induction with IPTG is a preferred method of expression. Bacteria are grown according to standard procedures in the art. Fresh or frozen bacteria cells are used for isolation of protein.

[00112] Alternatively, it is possible to purify PfKinI-1 from bacteria periplasm. After PfKinI-1 is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to those of skill in the art. To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM $MgSO_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

[00113] Suitable purification schemes for some specific kinesins are outlined in U.S. Serial Number 09/295,612, filed April 20, 1999, hereby expressly incorporated herein in its entirety for all purposes.

Standard Protein Separation Techniques For Purifying PfKinI-1

Solubility Fractionation

[00114] Often as an initial step, particularly if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol includes adding saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This concentration will precipitate the most hydrophobic of proteins. The precipitate is then discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, either through dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

Size Differential Filtration

[00115] The molecular weight of PfKinI-1 can be used to isolated it from proteins of greater and lesser size using ultrafiltration through membranes of different pore size (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

Column Chromatography

[00116] PfKinI-1 can also be separated from other proteins on the basis of its size, net surface charge, hydrophobicity, and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art. It will be apparent to one of skill that chromatographic

techniques can be performed at any scale and using equipment from many different manufacturers (e.g., Pharmacia Biotech).

Immunological Detection of PfKinI-1

[00117] In addition to the detection of PfKinI-1 genes and gene expression using nucleic acid hybridization technology, one can also use immunoassays to detect PfKinI-1. Immunoassays can be used to qualitatively or quantitatively analyze PfKinI-1. A general overview of the applicable technology can be found in Harlow & Lane, *Antibodies: A Laboratory Manual* (1988).

Antibodies to PfKinI-1

[00118] Methods of producing polyclonal and monoclonal antibodies that react specifically with PfKinI-1 are known to those of skill in the art (*see, e.g.*, Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (*see, e.g.*, Huse *et al.*, *Science* 246:1275-1281 (1989); Ward *et al.*, *Nature* 341:544-546 (1989)).

[00119] Humanized forms of mouse antibodies can be generated by linking the CDR regions of non-human antibodies to human constant regions by recombinant DNA techniques. See Queen et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:10029-10033 and PCT Publication WO 90/07861 (each of which is incorporated herein by reference for all purposes).

[00120] Human antibodies can be obtained using phage-display methods. See, e.g., Dower et al. PCT Publication WO 91/17271; McCafferty et al. PCT Publication WO 92/01047. In these methods, libraries of phage are produced in which members display different antibodies on their outer surfaces. Antibodies are usually displayed as Fv or Fab fragments. Phage displaying antibodies with a desired specificity are selected by affinity enrichment to PfKinI-1 or fragments thereof. Human antibodies against PfKinI-1 can also be produced from non-human transgenic mammals having transgenes encoding at least a segment of the human immunoglobulin locus and an inactivated endogenous immunoglobulin locus. See, e.g., Lonberg et al. PCT Publication WO 93/12227; Kucherlapati PCT Publication WO 91/10741 (each of which is incorporated herein by reference in its entirety for all purposes). Human antibodies can be selected by competitive binding experiments, or otherwise, to have the same

epitope specificity as a particular mouse antibody. Such antibodies are particularly likely to share the useful functional properties of the mouse antibodies. Human polyclonal antibodies can also be provided in the form of serum from humans immunized with an immunogenic agent. Optionally, such polyclonal antibodies can be concentrated by affinity purification using PfKinI-1 as an affinity reagent.

[00121] A number of PfKinI-1 comprising immunogens may be used to produce antibodies specifically reactive with PfKinI-1. For example, recombinant PfKinI-1 or an antigenic fragment thereof such as the motor domain, is isolated as described herein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells as described above, and purified as generally described above. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

[00122] Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (e.g., BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to PfKinI-1. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see* Harlow & Lane, *supra*).

[00123] Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*see* Kohler & Milstein, *Eur. J. Immunol.* 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a

monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse *et al.*, *Science* 246:1275-1281 (1989).

[00124] Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-PfKinI-1 proteins or even other homologous proteins from other organisms (e.g., *C. elegans* unc-104 or human Kif1A), using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better.

[00125] Once PfKinI-1 specific antibodies are available, PfKinI-1 can be detected by a variety of immunoassay methods. For a review of immunological and immunoassay procedures, see *Basic and Clinical Immunology* (Stites & Terr eds., 7th ed. 1991). Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay* (Maggio ed., 1980); and Harlow & Lane, *supra*.

Binding assays

[00126] Antibodies can be used for treatment or to identify the presence of PfKinI-1 having the sequence identity characteristics as described herein. Additionally, antibodies can be used to identify modulators of the interaction between the antibody and PfKinI-1 as further described below. While the following discussion is directed toward the use of antibodies in the use of binding assays, it is understood that the same general assay formats such as those described for "non-competitive" or "competitive" assays can be used with any compound which binds to PfKinI-1 such as microtubules or the compounds described in U.S. Patent No. 6,207,403.

[00127] In a preferred embodiment, PfKinI-1 is detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.,* U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology Volume 37: Antibodies in Cell Biology (*Asai, ed. 1993); *Basic and Clinical Immunology* (Stites & Terr, eds., 7th ed. 1991). Immunological binding assays (or immunoassays) typically use an antibody that specifically binds to a protein or antigen of choice (in this case the PfKinI-1 or antigenic subsequence thereof). The antibody (e.g., anti-PfKinI-1)

may be produced by any of a number of means well known to those of skill in the art and as described above.

[00128] Immunoassays also often use a labeling agent to specifically bind to and label the complex formed by the antibody and antigen. The labeling agent may itself be one of the moieties comprising the antibody/antigen complex. Thus, the labeling agent may be a labeled PfKinI-1 polypeptide or a labeled anti-PfKinI-1 antibody. Alternatively, the labeling agent may be a third moiety, such as a secondary antibody, that specifically binds to the antibody/PfKinI-1 complex (a secondary antibody is typically specific to antibodies of the species from which the first antibody is derived). Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G may also be used as the label agent. These proteins exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see generally* Kronval *et al.*, *J. Immunol.* 111:1401-1406 (1973); Akerstrom *et al.*, *J. Immunol.* 135:2589-2542 (1985)). The labeling agent can be modified with a detectable moiety, such as biotin, to which another molecule can specifically bind, such as streptavidin. A variety of detectable moieties are well known to those skilled in the art.

[00129] Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, preferably from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 4°C to 40°C.

Non-Competitive Assay Formats

[00130] Immunoassays for detecting PfKinI-1 in samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of antigen is directly measured. In one preferred "sandwich" assay, for example, the anti-PfKinI-1 antibodies can be bound directly to a solid substrate on which they are immobilized. These immobilized antibodies then capture PfKinI-1 present in the test sample. PfKinI-1 is thus immobilized is then bound by a labeling agent, such as a second PfKinI-1 antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibody is typically modified with a detectable moiety, such as biotin, to which another molecule specifically binds, e.g., streptavidin, to provide a detectable moiety.

Competitive assay formats

[00131] In competitive assays, the amount of PfKinI-1 present in the sample is measured indirectly by measuring the amount of a known, added (exogenous) PfKinI-1 displaced (competed away) from an anti-PfKinI-1 antibody by the unknown PfKinI-1 present in a sample. In one competitive assay, a known amount of PfKinI-1 is added to a sample and the sample is then contacted with an antibody that specifically binds to PfKinI-1. The amount of exogenous PfKinI-1 bound to the antibody is inversely proportional to the concentration of PfKinI-1 present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. The amount of PfKinI-1 bound to the antibody may be determined either by measuring the amount of PfKinI-1 present in a PfKinI-1/antibody complex, or alternatively by measuring the amount of remaining uncomplexed protein. The amount of PfKinI-1 may be detected by providing a labeled PfKinI-1 molecule.

[00132] A hapten inhibition assay is another preferred competitive assay. In this assay the known PfKinI-1, is immobilized on a solid substrate. A known amount of anti-PfKinI-1 antibody is added to the sample, and the sample is then contacted with the PfKinI-1. The amount of anti-PfKinI-1 antibody bound to the known immobilized PfKinI-1 is inversely proportional to the amount of PfKinI-1 present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

Cross-reactivity determinations

[00133] Immunoassays in the competitive binding format can also be used for crossreactivity determinations. For example, a protein at least partially encoded by SEQ ID NO:2 can be immobilized to a solid support. Proteins (e.g., *C. elegans* unc-104 or human Kif1A) are added to the assay that compete for binding of the antisera to the immobilized antigen. The ability of the added proteins to compete for binding of the antisera to the immobilized protein is compared to the ability of PfKinI-1 of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 to compete with itself. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the added proteins listed above are selected and pooled. The cross-reacting antibodies are optionally removed from the pooled

antisera by immunoabsorption with the added considered proteins, e.g., distantly related homologues.

[00134] The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps the protein of this invention, to the immunogen protein (i.e., PfKinI-1 of SEQ ID NO:2). In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required to inhibit 50% of binding is less than 10 times the amount of the protein of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 that is required to inhibit 50% of binding, then the second protein is said to specifically bind to the polyclonal antibodies generated to a PfKinI-1 immunogen.

Other assay formats

[00135] Western blot (immunoblot) analysis is used to detect and quantify the presence of PfKinI-1 in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support, (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter), and incubating the sample with the antibodies that specifically bind PfKinI-1. The anti-PfKinI-1 antibodies specifically bind to the PfKinI-1 on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (e.g., labeled sheep anti-mouse antibodies) that specifically bind to the anti-PfKinI-1 antibodies.

[00136] Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (e.g., antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see* Monroe *et al.*, *Amer. Clin. Prod. Rev.* 5:34-41 (1986)).

Reduction of non-specific binding

[00137] One of skill in the art will appreciate that it is often desirable to minimize non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of reducing such non-specific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous

composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used with powdered milk being most preferred.

Labels

[00138] The particular label or detectable group used in the assay is not a critical aspect of the invention, as long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels have been well-developed in the field of immunoassays and, in general, most any label useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (e.g., DYNABEADS^{TM}), fluorescent dyes (e.g., fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (e.g., ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (e.g., horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), colorimetric labels such as colloidal gold or colored glass or plastic beads (e.g., polystyrene, polypropylene, latex, etc.) or other labels that can be detected by mass spectroscopy, NMR spectroscopy, or other analytical means known in the art.

[00139] The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

[00140] Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (e.g., biotin) is covalently bound to the molecule. The ligand then binds to another molecules (e.g., streptavidin) molecule, which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. The ligands and their targets can be used in any suitable combination with antibodies that recognize PfKinI-1, or secondary antibodies that recognize anti-PfKinI-1.

[00141] The molecules can also be conjugated directly to signal generating compounds, e.g., by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, etc. Chemiluminescent compounds

include luciferin, and 2,3-dihydrophthalazinediones, e.g., luminol. For a review of various labeling or signal producing systems which may be used, see U.S. Patent No. 4,391,904.

[00142] Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, by means of photographic film, by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels may be detected simply by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

[00143] Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

Assays for modulators of the target protein

[00144] Modulators of PfKinI-1 (e.g., inhibitors) can be identified using assays such as those described in this section. Modulators identified by the assays can subsequently be formulated as pharmaceutical compositions for treatment of conditions associated with infection by *Plasmodium falciparum*. The substantial variation in the amino acid sequence of the motor domains from PfKinI-1 and its mammalian counterpart (approximately 60% similarity and 49% identity) enables modulators to be identified that have the appropriate species specificity.

Functional Assays

[00145] Assays that can be used to test for modulators of the target protein include a variety of *in vitro* or *in vivo* assays, e.g., microtubule gliding assays, binding assays such as microtubule binding assays, microtubule depolymerization assays, and ATPase assays (Kodama et al., *J. Biochem.* 99: 1465-1472 (1986); Stewart et al., *Proc. Nat'l Acad. Sci. USA* 90: 5209-5213 (1993); Lombillo et al., *J. Cell Biol.* 128:107-115 (1995); Vale et al., *Cell* 42:39-50 (1985)). Methods of performing motility assays are well known (see, e.g., Hall et al. (1996)

Biophys. J. 71:3467-3476; Turner et al. (1996) *Anal. Biochem.* 242:20-5; Gittes et al. (1996) *Biophys. J.* 70:418-29; Shirakawa et al. (1995) *J. Exp. Biol.* 198:1809-15; Winkelmann et al. (1995) *Biophys J.* 68:2444-53; Winkelmann et al. (1995) *Biophys. J.* 68:72S, and the like).

[00146] Modulation is tested by screening for candidate agents capable of modulating the activity of the target protein comprising the steps of combining a candidate agent with the target protein, as above, and determining an alteration in the biological activity of the target protein. Thus, in this embodiment, the candidate agent should both bind to the target protein (although this may not be necessary), and alter its biological or biochemical activity as defined herein. The methods include both *in vitro* screening methods and *in vivo* screening of cells for alterations in cell cycle distribution, cell viability, or for the presence, morphology, activity, distribution, or amount of mitotic spindles, as are generally outlined above.

[00147] A preferred assay for high throughput screening is an ATPase assay with colorimetric detection, e.g., malachite green for end-point detection or coupled PK/LDH for continuous rate monitoring. An exemplary ATPase activity assay utilizes 0.3M PCA (perchloric acid) and malachite green reagent (8.27 mM sodium molybdate II, 0.33 mM malachite green oxalate, and 0.8 mM triton X-100). To perform the assay, 10 μL of reaction is quenched in 90 μL of cold 0.3 M PCA. Phosphate standards are used so data can be converted to mM inorganic phosphate released. When all reactions and standards have been quenched in PCA, 100 μL of malachite green reagent is added to the relevant wells in e.g., a microtiter plate. The mixture is developed for 10-15 minutes and the plate is read at an absorbance of 650 nm. If phosphate standards were used, absorbance readings can be converted to mM Pi and plotted over time. Additionally, ATPase assays known in the art include the luciferase assay.

[00148] Another exemplary assay can be performed using the following two specific solutions. Solution A contains 1 mM ATP, 2 mM phosphoenolpyruvate in a working buffer (25 mM Pipes pH 6.8), 2 mM $MgCl_2$, 1 mM EGTA, 1 mM DTT, 5 μM Taxol™, 25 ppm antifoam, pH 6.8. Solution B contains 0.6 mM NADH, 0.2 mg/ml BSA, 1:100 dilution of PK/LDH mixture from Sigma, 200 μg/ml microtubules, 100 nM PfKinI-1 (i.e., about 2.5 μg/ml). To initiate the experiment, 1 μl of DMSO stock of test compound is added to each well of the bottom row of a 96-well half area plate. Control wells contain only DMSO alone. 50 μl of solution A is then added to each well. The solutions are mixed by repeated pipetting, followed by a series of dilutions by repeated transferring of 50 μl of solution between rows. The reaction is initiated by adding 50 μl of solution B. The plate is then inserted in the reader and absorbance at 340 nm is monitored for five minutes. The observed rate for 50 μL Solution A +

50 μL Solution B in a half-area plate should be about 100 mOD/min. Optionally, a series of dilutions is made and absorbance similarly measured. Similar procedures can be used to study the inhibitory effect of a test agent on the basal (i.e., not microtubule-dependent) ATPase of PfKinI-1. In these assays, microtubules are omitted from Solution B and PfKinI-1 concentration is increased to at least 2 mM.

[00149] Such assays can be used to test for the activity of PfKinI-1 isolated from endogenous sources or recombinant sources. Furthermore, such assays can be used to test for modulators of PfKinI-1. Modulators can increase or decrease activity of PfKinI-1.

[00150] In a preferred embodiment, molecular motor activity is measured by the methods disclosed in Serial No. 09/314,464, filed May 18, 1999, entitled "Compositions and assay utilizing ADP or phosphate for detecting protein modulators", which is incorporated herein by reference in its entirety. More specifically, this assay detects modulators of any aspect of a kinesin motor function ranging from interaction with microtubules to hydrolysis of ATP. ADP or phosphate is used as the readout for protein activity.

[00151] There are a number of enzymatic assays known in the art which use ADP as a substrate. For example, kinase reactions such as pyruvate kinases are known. See, *Nature* 78:632 (1956) and *Mol. Pharmacol.* 6:31 (1970). This is a preferred method in that it allows the regeneration of ATP. In one embodiment, the level of activity of the enzymatic reaction is determined directly. In a preferred embodiment, the level of activity of the enzymatic reaction which uses ADP as a substrate is measured indirectly by being coupled to another reaction. For example, in one embodiment, the method further comprises a lactate dehydrogenase reaction under conditions which normally allow the oxidation of NADH, wherein said lactate dehydrogenase reaction is dependent on the pyruvate kinase reaction. Measurement of enzymatic reactions by coupling is known in the art. Furthermore, there are a number of reactions which utilize phosphate. Examples of such reactions include a purine nucleoside phosphorylase reaction. This reaction can be measured directly or indirectly. A particularly preferred embodiment utilizes the pyruvate kinase/lactate dehydrogenase system.

[00152] In one embodiment, the detection of the ADP or phosphate proceeds non-enzymatically, for example, by binding or reacting the ADP or phosphate with a detectable compound. For example, phosphomolybdate based assays may be used which involve conversion of free phosphate to a phosphomolybdate complex. One method of quantifying the phosphomolybdate is with malachite green. Alternatively, a fluorescently labeled form of a

phosphate binding protein, such as the *E. coli* phosphate binding protein, can be used to measure phosphate by a shift in its fluorescence.

[00153] In addition, target protein activity can be examined by determining modulation of target protein in vitro using cultured cells. The cells are treated with a candidate agent and the effect of such agent on the cells is then determined either directly or by examining relevant surrogate markers. For example, characteristics such as mitotic spindle morphology and cell cycle distribution can be used to determine the effect.

[00154] Thus, in a preferred embodiment, the methods comprise combining a target protein and a candidate agent, and determining the effect of the candidate agent on the target protein. Generally a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a differential response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration or below the level of detection.

[00155] As will be appreciated by those of skill in the art, the components may be added in buffers and reagents to assay target protein activity and give optimal signals. Since the methods allow kinetic measurements, the incubation periods can be optimized to give adequate detection signals over the background.

[00156] In a preferred embodiment, an antifoam or a surfactant is included in the assay mixture. Suitable antifoams include, but are not limited to, antifoam 289 (Sigma). Suitable surfactants include, but are not limited to, Tween, Tritons, including Triton X-100, saponins, and polyoxyethylene ethers. Generally, the antifoams, detergents, or surfactants are added at a range from about 0.01 ppm to about 10 ppm.

[00157] A preferred assay design is also provided. In one aspect, the invention provides a multi-time-point (kinetic) assay, with at least two data points being preferred. In the case of multiple measurements, the absolute rate of the protein activity can be determined.

Binding Assays

[00158] In a preferred embodiment, the binding of the candidate agent is determined through the use of competitive binding assays. In this embodiment, the competitor is a binding moiety known to bind to the target protein, such as an antibody, peptide, binding partner, ligand, etc. Under certain circumstances, there may be competitive binding as between the candidate agent and the binding moiety, with the binding moiety displacing the candidate agent.

[00159] Competitive screening assays may be done by combining the target protein and a drug candidate in a first sample. A second sample comprises a candidate agent, the target

protein and a compound that is known to modulate the target protein. This may be performed in either the presence or absence of microtubules. The binding of the candidate agent is determined for both samples, and a change, or difference in binding between the two samples indicates the presence of an agent capable of binding to the target protein and potentially modulating its activity. That is, if the binding of the candidate agent is different in the second sample relative to the first sample, the candidate agent is capable of binding to the target protein.

[00160] In one embodiment, the candidate agent is labeled. Either the candidate agent, or the competitor, or both, is added first to the target protein for a time sufficient to allow binding. Incubations may be performed at any temperature which facilitates optimal activity, typically between 4 and 40°C. Incubation periods are selected for optimum activity, but may also be optimized to facilitate rapid high throughput screening. Typically between 0.1 and 1 hour will be sufficient. Excess reagent is generally removed or washed away. The second component is then added, and the presence or absence of the labeled component is followed, to indicate binding.

[00161] In a preferred embodiment, the competitor is added first, followed by the candidate agent. Displacement of the competitor is an indication the candidate agent is binding to the target protein and thus is capable of binding to, and potentially modulating, the activity of the target protein. In this embodiment, either component can be labeled. Thus, for example, if the competitor is labeled, the presence of label in the wash solution indicates displacement by the agent. Alternatively, if the candidate agent is labeled, the presence of the label on the support indicates displacement.

[00162] In an alternative embodiment, the candidate agent is added first, with incubation and washing, followed by the competitor. The absence of binding by the competitor may indicate the candidate agent is bound to the target protein with a higher affinity. Thus, if the candidate agent is labeled, the presence of the label on the support, coupled with a lack of competitor binding, may indicate the candidate agent is capable of binding to the target protein.

Microtubule depolymerization assay protocol

[00163] Microtubule depolymerization assays generally involve contacting a polymerized microtubule with a microtubule severing protein or a microtubule depolymerizing protein (i.e., PfKinI-1) in the presence of a chemical energy source (e.g., ATP or GTP), and the candidate agent; and detecting and/or quantifying the formation of microtubule degradation products (e.g., tubulin monomers). Agents that inhibit the activity of the microtubule depolymerization or

severing proteins will inhibit the breakdown of the polymerized microtubules thereby delaying the formation of or reducing the quantity of tubulin monomers or oligomers. Thus, a decrease in the rate of formation or amount of tubulin monomer or an increase in the ratio of tubulin polymer (microtubules) to tubulin monomer indicates an inhibitory modulating effect of the agent. Conversely, an increase in the rate of formation or amount of tubulin monomer or an increase in the ratio of tubulin polymer (microtubules) to tubulin monomer indicates a microtubule stabilizing effect of the agent. See, e.g., Desai et al. (1999) *Cell* 96(1):69-78; and Hartman et al. (1998) *Cell* 93(2):277-87, each of which is incorporated herein by reference for all purposes.

[00164] The increase or decrease is determined by reference to one or more controls. A control is essentially an identical assay that either lacks the test agent or contains a "reference" agent that has a known activity. Assays lacking any test agent whatsoever act as negative controls, while assays utilizing an agent that has known modulatory activity act as positive controls.

[00165] In a preferred embodiment, the assay is scored as positive for a stabilizing agent when there is a significant difference between the negative control and the test assay and/or when there is no significant difference between the positive control and the test assay. The significant difference is preferably a statistically significant difference, more preferably at least about a 10% difference, and most preferably at least about a 20%, 30%, 50%, or 100% difference.

[00166] The assays can be performed in solution or in the solid phase (i.e., with one or more components of the assay attached to a solid surface). See, e.g., PCT Publication, WO 99/53295, which is incorporated herein by reference in its entirety.

Candidate agents

[00167] Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 100 and less than about 2,500 daltons. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among biomolecules including

peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogs or combinations thereof. Particularly preferred are peptides.

[00168] Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. In a preferred embodiment, the candidate agents are organic chemical moieties, a wide variety of which are available in the literature. Combinatorial libraries can be produced for many types of compounds that can be synthesized in step-by-step fashion. Such compounds include polypeptides, proteins, nucleic acids, beta-turn mimetics, polysaccharides, phospholipids, hormones, prostaglandins, steroids, aromatic compounds, heterocyclic compounds, benzodiazepines, oligomeric N-substituted glycines and oligocarbamates. Large combinatorial libraries of compounds can be constructed using the methods described in PCT Publication No. WO 95/12608, WO 93/06121, 94/08051, WO 95/35503 and WO 95/30642 (each of which is incorporated herein by reference in its entirety for all purposes). Peptide libraries can also be generated by phage display methods. See, PCT Publication No. WO 91/18980. Compounds to be screened can also be obtained from governmental or private sources, including, e.g., the National Cancer Institute's (NCI) Natural Product Repository, Bethesda, MD; the NCI Open Synthetic Compound Collection, Bethesda, MD; NCI's Developmental Therapeutics Program, and the like.

Other assay components

[00169] The assays provided utilize target protein as defined herein. In one embodiment, portions of target protein are utilized; in a preferred embodiment, portions having target protein activity as described herein are used. In addition, the assays described herein may utilize either isolated target proteins or cells or animal models comprising the target proteins.

[00170] A variety of other reagents may be included in the screening assays. These include reagents like salts, neutral proteins, e.g. albumin, detergents, etc which may be used to facilitate optimal protein-protein binding and/or reduce non-specific or background interactions. Also, reagents that otherwise improve the efficiency of the assay, such as protease inhibitors, nuclease inhibitors, anti-microbial agents, etc., may be used. The mixture of components may be added in any order that provides for the requisite binding.

Optimization of Lead Compounds

[00171] Modulators identified as having activity based upon assays such as those just described can be further evaluated and modified to increase their activity. Compounds that inhibit the ATPase and/or microtubule depolymerization activity of pfKinI-1 for instance can be

further tested for antiproliferative activity toward drug-sensitive and drug-resistant parasites grown in culture. These compounds can also be evaluated for their toxicity towards mammalian cells in culture.

[00172] Those compounds exhibiting good activity with minimal toxicity can then optionally serve as lead compounds for the synthesis of analog compounds that exhibit even higher activity (i.e., increased inhibition of ATPase or depolymerization activity). Such analogs are typically synthesized to have an electronic configuration and a molecular conformation similar to that of the lead compound. Identification of analog compounds can be performed through use of techniques such as self-consistent field (SCF) analysis, configuration interaction (CI) analysis, and normal mode dynamics analysis. Computer programs for implementing these techniques are available. See, e.g., Rein et al., (1989) Computer-Assisted Modeling of Receptor-Ligand Interactions (Alan Liss, New York).

[00173] Once analogs have been prepared, they can be screened using the methods disclosed supra to identify those analogs that exhibit an increase in the desired activity. Such compounds can then be subjected to further analysis to identify those compounds that appear to have the greatest potential as pharmaceutical agents. Alternatively, analogs shown to have activity through the screening methods can themselves serve as lead compounds in the preparation of still further analogs, which can be screened by the methods disclosed herein. The cycle of screening, synthesizing analogs and rescreening can be repeated multiple times to further optimize the activity of the analog.

[00174] Further guidance on the synthesis of analog compounds and lead optimization is provided by, for example: Iwata, Y., et al. (2001) *J. Med. Chem.* 44:1718-1728; Prokai, L., et al. (2001) *J. Med. Chem.* 44:1623-1626; Roussel, P. et al., (1999) *Tetrahedron* 55:6219-6230; Bunin, B.A., et al. (1999) *Ann. Rep. Med. Chem.* 34:267-286; Venkatesh, S., et al. (2000) *J. Pharm. Sci.* 89:145-154; and Bajpai, M. and Adkinson, K.K. (2000) *Curr. Opin. Drug Discovery and Dev.* 3:63-71.

Applications

[00175] The methods of the invention are used to identify compounds useful in either the treatment of an existing infection by the malaria parasite or in the prevention of such an infection from occurring in the first place.

[00176] Accordingly, the compositions of the invention are administered to cells. By "administered" herein is meant administration of a therapeutically effective dose of the candidate agents of the invention to a cell either in cell culture or in a patient. By

"therapeutically effective dose" herein is meant a dose that produces the effects for which it is administered. The exact dose will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques. As is known in the art, adjustments for systemic versus localized delivery, age, body weight, general health, sex, diet, time of administration, drug interaction and the severity of the condition may be necessary, and will be ascertainable with routine experimentation by those skilled in the art. By "cells" herein is meant almost any cell in which mitosis or meiosis can be altered.

[00177] A "patient" for the purposes of the present invention includes both humans and other animals, particularly mammals, and other organisms. Thus the methods are applicable to both human therapy and veterinary applications. In the preferred embodiment the patient is a mammal, and in the most preferred embodiment the patient is human.

[00178] Candidate agents having the desired pharmacological activity may be administered in a physiologically acceptable carrier to a patient, as described herein. Depending upon the manner of introduction, the compounds may be formulated in a variety of ways as discussed below. The concentration of therapeutically active compound in the formulation may vary from about 0.1-100 wt.%. The agents maybe administered alone or in combination with other treatments, i.e., other antimalarial agents.

[00179] In a preferred embodiment, the pharmaceutical compositions are in a water soluble form, such as pharmaceutically acceptable salts, which is meant to include both acid and base addition salts.

[00180] The pharmaceutical compositions can be prepared in various forms, such as granules, tablets, pills, suppositories, capsules, suspensions, salves, lotions and the like. Pharmaceutical grade organic or inorganic carriers and/or diluents suitable for oral and topical use can be used to make up compositions containing the therapeutically-active compounds. Diluents known to the art include aqueous media, vegetable and animal oils and fats. Stabilizing agents, wetting and emulsifying agents, salts for varying the osmotic pressure or buffers for securing an adequate pH value, and skin penetration enhancers can be used as auxiliary agents. The pharmaceutical compositions may also include one or more of the following: carrier proteins such as serum albumin; buffers; fillers such as microcrystalline cellulose, lactose, corn and other starches; binding agents; sweeteners and other flavoring agents; coloring agents; and polyethylene glycol. Additives are well known in the art, and are used in a variety of formulations.

[00181] The administration of the candidate agents of the present invention can be done in a variety of ways as discussed above, including, but not limited to, orally, subcutaneously,

intravenously, intranasally, transdermally, intraperitoneally, intramuscularly, intrapulmonary, vaginally, rectally, or intraocularly. In some instances, for example, in the treatment of wounds and inflammation, the candidate agents may be directly applied as a solution or spray.

[00182] One of skill in the art will readily appreciate that the methods described herein also can be used for diagnostic applications. A diagnostic as used herein is a compound or method that assists in the identification and characterization of a health or disease state in humans or other animals. More specifically, antibodies which specifically bind PfKinI-1 may be used for the diagnosis of disorders characterized by expression of PfKinI-1 or in assays to monitor patients being treated with PfKinI-1, or agonists, antagonists, or inhibitors of PfKinI-1. Diagnostic assays include methods which utilize the antibody and a label to detect PfKinI-1 in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecules. A wide variety of reporter molecules are known in the art and may be used.

[00183] In another embodiment of the invention, the polynucleotides encoding PfKinI-1 may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantitate gene expression in biopsied tissues in which PfKinI-1 may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of PfKinI-1, and to monitor regulation of PfKinI-1 levels during therapeutic intervention.

[00184] In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences encoding PfKinI-1 or closely related molecules may be used. The specificity of the probe, whether it's made from a highly specific region or from a less specific region, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low) will determine whether the probe identifies only naturally occurring sequences encoding PfKinI-1, allelic variants, or related sequences.

[00185] Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the PfKinI-1 encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11 or from genomic sequences including promoters, enhancers, and introns of the PfKinI-1 gene.

[00186] Means for producing specific hybridization probes for DNAs encoding PfKinI-1 include the cloning of polynucleotide sequences encoding PfKinI-1 or derivatives thereof into

vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems and the like.

[00187] In a particular aspect, the nucleotide sequences encoding PfKinI-1 may be useful in assays that detect the presence of associated disorders. The nucleotide sequences encoding PfKinI-1 may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantitated and compared with a standard value or baseline value. If the amount of signal is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding PfKinI-1 in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

[00188] Additional diagnostic uses for oligonucleotides designed from the sequences encoding PfKinI-1 may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding PfKinI-1, or a fragment of a polynucleotide complementary to the polynucleotide encoding PfKinI-1, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantitation of closely related DNA or RNA sequences.

[00189] Methods which may be used to quantitate the expression of PfKinI-1 include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. The speed of quantitation of multiple samples may be accelerated by running the assay in an ELISA format where the oligomer of interest is presented in various dilutions and a spectrophotometer or colorimetric response gives rapid quantitation.

[00190] One of skill in the art will readily appreciate that the methods described herein also can be used for diagnostic applications. A diagnostic as used herein is a compound or method that assists in the identification and characterization of a health or disease state in humans or other animals.

[00191] The present invention also provides for kits for screening for modulators of the target protein. Such kits can be prepared from readily available materials and reagents. For example, such kits can comprise any one or more of the following materials: biologically active target protein, reaction tubes, and instructions for testing activity of the target protein. Preferably, the kit contains biologically active target protein. A wide variety of kits and components can be prepared according to the present invention, depending upon the intended user of the kit and the particular needs of the user. For example, the kit can be tailored for ATPase assays, microtubule gliding assays, or microtubule binding assays.

[00192] The kinesins of the present invention and, in particular, their motor domains can be used for separation of a specific ligand from a heterologous mixture in aqueous solution as described by Stewart, U.S. Patent No. 5,830,659. In the system discussed by Stewart, a kinesin motor domain is linked to a ligand binding moiety, such as streptavidin. The chimeric kinesin motor domains are placed into a loading chamber containing the heterogeneous mixtures which is coupled to a receiving chamber by a channel bearing immobilized, aligned microtubules. Addition of ATP to the loading chamber results in translocation of the kinesin motor domains, now attached covalently to the receiving chamber. Hence, the ATP-driven motility activity of the kinesin motor domain results in separation of the desired ligand from the heterogeneous mixture. Stewart further teaches that all kinesin motor domains are suitable for use in the separation system.

[00193] The kinesins of their invention and, in particular, their motor domains can also be used in the field of nanotechnology. Molecular motors such as kinesin have widespread application in the construction of nanoscale machines. Biomolecular motors have real-world application in the emerging nanotechnological arts. For example, a 1999 NASA study identifies multiple applications for nanoscale motors – and kinesin in particular- in the aerospace field. Kinesin motor domains can be used in the construction of rotors and other mechanical components (for review, see Limberis and Stewart (2000) *Nanotechnology* 11:47-51) as well as light-operated molecular shuttles useful for nanoscale switches and pumps.

[00194] Nucleic acids encoding the kinesins of the invention are also useful for inclusion on a GeneChip™ array or the like for use in expression monitoring (see US 6,040,138, EP 853,679 and WO 97/27317). Such arrays typically contain oligonucleotide or cDNA probes to allow detection of large numbers of mRNAs within a mixture. Many of the nucleic acids included in such arrays are from genes or ESTs that have not been well characterized. Such arrays are often used to compare expression profiles between different tissues or between different conditions of the same tissue (healthy vs. diseased or drug-treated vs. control) to

identify differentially expressed transcripts. The differentially expressed transcripts are then useful, e.g., for diagnosis of disease states, or to characterize responses of drugs. The nucleic acids of the invention can be included on GeneChip™ arrays or the like together with probes containing a variety of other genes. The present nucleic acids are particularly useful for inclusion in GeneChip™ arrays for analyzing the cell cycle or proliferation state of cells. Nucleic acids encoding PfKinI-1 can be combined with nucleic acids encoding other kinesin molecules and/or nucleic acids from other genes having roles in DNA replication, cell division or other cell cycle function. Such arrays are also useful for analyzing candidate drugs for roles in modulation of the cell cycle and proliferation. The efficacy of such drugs can be assayed by determining the effect of the drug on the expression profile of genes affecting proliferation and the cell cycle.

Examples

Assay Protocol

[00195] This assay is based on detection of ADP production from a target protein's microtubule stimulated ATPase. ATP production is monitored by a coupled enzyme system consisting of pyruvate kinase and lactate dehydrogenase. Under the assay conditions described below, pyruvate kinase catalyzes the conversion of ADP and phosphoenol pyruvate to pyruvate and ATP. Lactate dehydrogenase then catalyzes the oxidation-reduction reaction of pyruvate and NADH to lactate and NAD+. Thus, for each molecule of ADP produced, one molecule of NADH is consumed. The amount of NADH in the assay solution is monitored by measuring light absorbance at a wavelength of 340 nm.

[00196] The final 25 µl assay solution consists of the following: 5 µg/ml target protein, 30 µg/ml microtubules, 5 µM Taxol™, 0.8 mM NADH, 1.5 mM phosphoenol pyruvate, 3.5 U/ml pyruvate kinase, 5 U/ml lactate dehydrogenase, 25 mM Pipes/KOH pH 6.8, 2mM $MgCl_2$, 1 mM EGTA, 1 mM MDTT, 0.1 mg/ml BSA, 0.001% antifoam 289, and 1 mM ATP.

[00197] Potential candidate agents are dissolved in DMSO at a concentration of about 1 mg/ml and 0.5 µl of each chemical solution is dispensed into a single well of a clear 384 well plate. Each of the 384 wells are then filled with 20 µl of a solution consisting of all of the assay components described above except for ATP. The plate is agitated at a high frequency. To start the assay, 5 µl of a solution containing ATP is added to each well. The plate is agitated and the absorbance is read at 340 nm over various time intervals. The assay is run at room temperature.

[00198] The assay components and the performance of the assay are optimized together to match the overall read time with the rate of the target protein's ADP production. The read time should be long enough for the rate of NADH consumption to reach steady state beyond an initial lag time of several seconds.

ATPase assay protocol

[00199] 2 μg/ml of the target protein is assayed for its microtubule-stimulated ATPase activity in a reaction buffer composed of 40 μM MES/KOH pH 6.8, 2 mM $MgCl_2$, 1 mM DTT, 1 mM EGTA, 100 μM ATP, 10 μM paclitaxel, 0.1 mg/ml BSA, 0.5 mM NADH, 1.5 mM phosphoenolpyruvate, lactate dehydrogenase/pyruvate kinase mix (Sigma, diluted 1:200 v/v final) and varied amounts of microtubules (from 50 μg/ml to 24 ng/ml). The reaction progress is followed over time by monitoring absorbance of the reaction mixture at 340 nm using a microtiter plate reader (SpectraMAX340, Molecular Devices, Inc.). The rate of the absorbance change is converted to μM NADH oxidized per second by referencing it to a set of standard NADH solutions of known concentrations. In this coupled ATPase assay, conversion of one NADH to NAD+ reports appearance of one ADP molecule, and as such, reports a single turnover of the target molecule ATPase.

[00200] Using this assay, kinetic parameters for the hydrolysis of ATP by PfKinI-1 have been calculated. A plot of the kinetic data are shown in FIG. 1A; reciprocal plot analysis shows that k_{cat} =1.8 sec^{-1} and that K_m (microtubules) = 0.13 μM.

Cloning of the *Plasmodium falciparum* KinI motor construct

[00201] *Plasmodium falciparum* Dd2 mixed stage cDNA library was obtained through the Malaria Research and Reference Reagent Resource Center (MR4), Division of Microbiology and Infectious Diseases, NIAID, NIH. A KinI-like region was recognized in the genomic sequence from the *Plasmodium* sequencing project (http://plasmodb.org; The *Plasmodium* Genome Database Collaborative, 2001). This sequence was used to design two primers (sense primer: 5'-ATGATAAAAGTTGTGGTGAGGAAG-3' (SEQ ID NO:13), antisense primer: 5'-GTTCTTAACTCTTGAAGAATATCTTAATG-3' (SEQ ID NO:14)) and together with pfu turbo DNA polymerase (Stratagene, La Jolla, CA), these were used to amplify the KinI motor domain fragment (from Isoleucine2 to Asparagine330; SEQ ID NO:11 for oligonucleotide and SEQ ID NO:12 for encoded protein) from the cDNA library. The PCR fragment was treated with Taq DNA polymerase to add A overhangs and then cloned into the

expression vector pCRT7/CT TOPO according to the manufacturer instructions (Invitrogen, Carlsbad, CA).

PfKinI-1 expression and purification

[00202] *E. coli* cells (BL21 Star (DE3) from Invitrogen Inc.) carrying PfKinI-1 expression vector were grown in LB medium supplemented with 100ug/ml ampicillin. Erlenmeyer flasks containing 1L of the medium were inoculated with bacterial colonies from a freshly transformed cells and grown for 24h at room temperature. No additional induction was required since there was a substantial basal level of expression observed.

[00203] The cells were harvested by centrifugation and resuspended in a lysis buffer (50 mM TRIS/HCl; 50 mM KCl; 10 mM Imidazole; 2 mM $MgCl_2$;1mM ATP;1mM PMSF; 8 mM β-mercaptoethanol; pH 7.4). Cells were disrupted by several passages through microfluidizer and the insoluble bacterial debris separated by centrifugation (45min, 45000rpm, 45Ti rotor). The soluble fraction was applied to the Ni-NTA agarose column, washed with two column volumes of the lysis buffer and eluted in the lysis buffer supplemented with 300mM Imidazole. Protein containing fractions were pooled and dialyzed overnight against lysis buffer without imidazole. The dialyzed protein was bound to the SP-sepharose column and eluted with a gradient of KCl.

[00204] Typically, this procedure produced 20 mg of a soluble protein, with over 95% purity as judged by coomassie stained SDS/PAGE, per 1L of bacterial culture.

Microtubule depolymerization assay

[00205] For the microtubule depolymerization assay, 100 µl reaction mixes containing the PfKinI-1 construct, Taxol™-stabilized microtubules and adenine nucleotide were assembled. ATP,ADP, AMPPNP and Apyrase (for no nucleotide-free state) were the typical nucleotide conditions tested in this assay. Two control mixes were routinely included from which PfKinI-1 or microtubules were omitted. Samples were incubated at room temperature for 15min. Following the incubation, samples were centrifuged at 100,000g for 10 min. Supernatants were aspirated away from the microtubule pellets and mixed with an equal volume of 2x SDSPAGE loading dye. Pellets were resuspended in a volume of 1x loading dye equivalent to 2x starting volume of the sample before spin. Both supernatant and pellet samples were loaded on an SDS PAGE gel, run and stained with Coomassie Blue according to standard procedures. Appearance of tubulin (about 50 kDa) in the lanes containing the supernatant was

interpreted as a sign of microtubule depolymerization. An example of the results from such an experiment in which PfKinI-1 was included or omitted (control) are shown in FIG. 1B.

[00206] Other microtubule depolymerization assays take advantage of the fact that certain fluorescence dyes have higher fluorescence intensity when bound to polymerized versus free tubulin. These assays generally involve contacting a sample containing polymerized microtubules with target protein (e.g., PfKinI-1) in the presence of ATP and a fluorescent dye. Fluorescence intensity is then measured as a function of time, with a decrease in intensity over time reflecting the conversion of microtubules to tubulin. A variety of dyes can be utilized in these assays including, for example, 4',6'-diamidino-2-phenylindole (DAPI), anilinonapthalene sulfonate (ANS), bis-ANS, N-phenyl-1-naphthylene (NPN), DCVJ (each of the foregoing available from Molecular Probes, OR), ruthenium red and cresol violet. Further guidance regarding such assays is provided in PCT Publication WO 99/53295 and by Bonne et al. (1985) *J. Biol. Chem.* 260:2819-2825. Using these assays, microtubule disassembly can be monitored in real time; the assays are also suitable for high throughput screening of candidate agents to determine if they function as modulators of PfKinI-1 depolymerization activity. The assays, for example, can be performed in a 384-well format.

[00207] Results from an assay containing PfKinI-1, polymerized microtubules and DAPI are shown in FIG. 2A. The decrease in fluorescence as a function of time provides further evidence that PfKinI-1 has depolymerization activity. FIG. 2B shows the slope values corresponding to the rate of depolymerization reactions catalyzed by PfKinI-1as compared to controls lacking either enzyme or ATP.

[00208] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, and patent applications cited herein are hereby incorporated by reference in their entirety.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid sequence encoding a protein, wherein the protein has the following properties: (i) the protein has microtubule-stimulated ATPase activity; and (ii) the protein comprises a sequence that has greater than 90% amino acid sequence identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12, as measured using a sequence comparison algorithm.

2. An isolated nucleic acid sequence of claim 1, wherein the protein specifically binds to polyclonal antibodies generated against a protein comprising SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

3. An isolated nucleic acid sequence of claim 1, wherein the nucleic acid encodes SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

4. An isolated nucleic acid sequence of claim 1, wherein the nucleic acid comprises a polynucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11.

5. An isolated nucleic acid sequence of claim 1, wherein the nucleic acid hybridizes under stringent hybridization conditions to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11.

6. An expression vector comprising a nucleic acid encoding a protein, wherein the protein has the following properties: (i) the protein has microtubule stimulated ATPase activity; and (ii) the protein comprises a sequence that has greater than 90% amino acid sequence identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 as measured using a sequence comparison algorithm.

7. A host cell transfected with the vector of claim 6.

8. An isolated protein, wherein the protein comprises a sequence that has greater than 90% amino acid sequence identity to SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12 as measured using a sequence comparison algorithm and wherein the protein has microtubule stimulated ATPase activity.

9. An isolated protein of claim 8, wherein the protein specifically binds to polyclonal antibodies generated against a protein comprising SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

10. An isolated protein of claim 8, wherein the protein comprises SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12.

11. An isolated protein comprising an amino acid sequence of SEQ ID NO:2.

12. An isolated protein comprising an amino acid sequence of SEQ ID NO:4.

13. An isolated protein comprising an amino acid sequence of SEQ ID NO:6.

14. An isolated protein comprising an amino acid sequence of SEQ ID NO:8.

15. An isolated protein comprising an amino acid sequence of SEQ ID NO:12.

16. An isolated nucleic acid comprising a sequence which has greater than 90% sequence identity with nucleotide SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, or SEQ ID NO:11 and which encodes a protein having microtubule-stimulated ATPase activity.

17. A method for screening a compound for anti-malarial activity, which method comprises

contacting the compound with a protein comprising SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, or SEQ ID NO:12; and

determining whether the compound binds to and inhibits the protein, any such binding and inhibition suggesting that the compound may have anti-malarial activity.

18. A method of claim 17, wherein the screening occurs in a multi-well plate as part of a high-throughput screen.

19. An antibody that specifically binds to a protein having a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10; and SEQ ID NO:12.

20. The antibody of claim 19 that is a polyclonal antibody.

21. The antibody of claim 19 that is a monoclonal antibody.

22. The antibody of claim 19 that is a recombinant antibody or an antibody fragment.

23. A pharmaceutical composition comprising an agent that inhibits the microtubule-stimulated ATPase and/or microtubule depolymerization activity of a protein having a sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10; and SEQ ID NO:12 and a pharmaceutically acceptable carrier.

24. The pharmaceutical composition of claim 23, wherein the agent is an antibody.

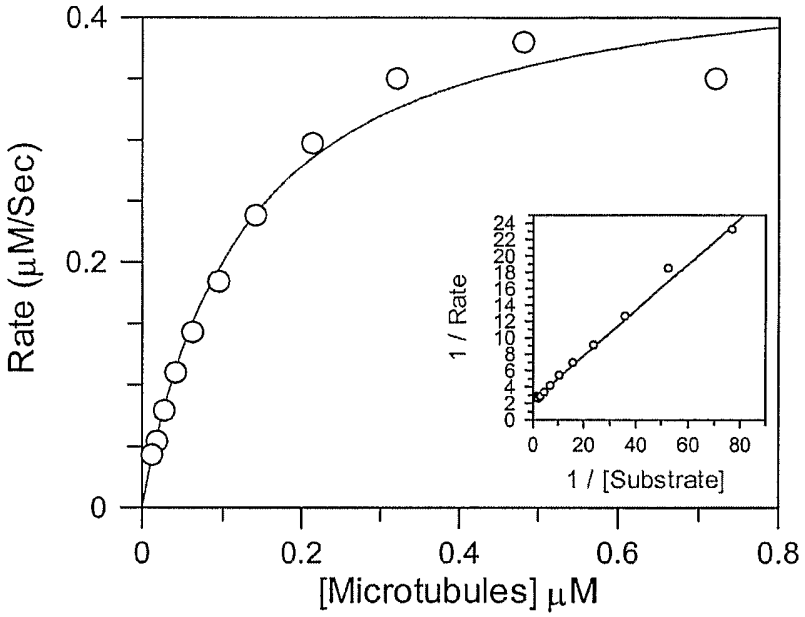

0.4
0.2
0
Rate (μM/Sec)
0
0.2
0.4
0.6
0.8
[Microtubules] μM
1 / Rate
0
2
4
6
8
10
12
14
16
18
20
22
24
0
20
40
60
80
1 / [Substrate]

FIG. 1A

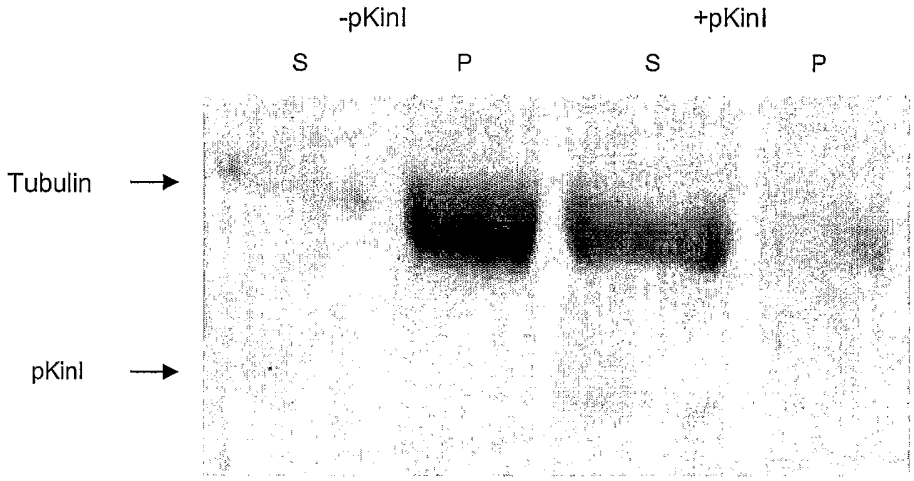

-pKinI
+pKinI
S
P
S
P
Tubulin
pKinI

FIG. 1B

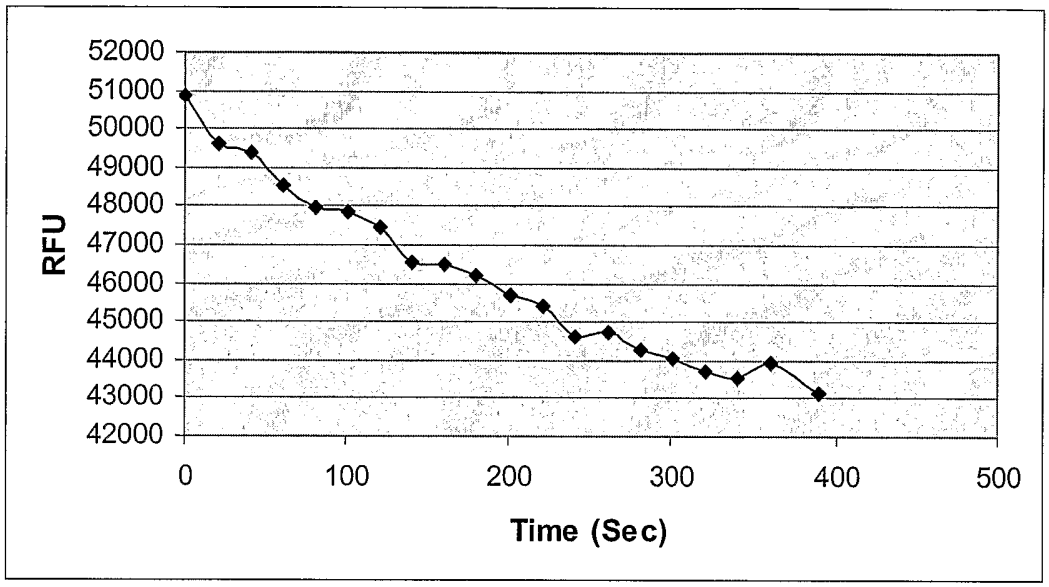

52000
51000
50000
49000
48000
47000
46000
45000
44000
43000
42000
RFU
0
100
200
300
400
500
Time (Sec)

FIG. 2A

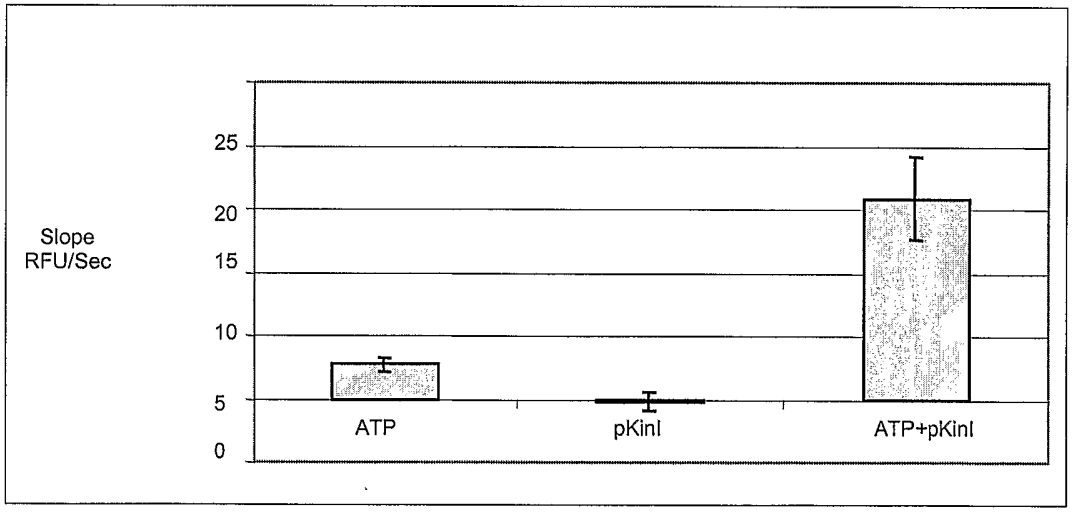

Slope
RFU/Sec
25
20
15
10
5
0
ATP
pKinI
ATP+pKinI

FIG. 2B

SEQUENCE LISTING

<110> Sakowicz, Roman
Beraud, Christophe
Guo, Jun
Freedman, Richard
Cytokinetics, Inc.

<120> Novel Motor Protein of P. falciparum and Methods for Its Use

<130> 020552-007210PC

<140> WO Not yet assigned
<141> Not yet assigned

<150> US 10/006,780
<151> 2001-11-30

<150> US 10/086,935
<151> 2002-02-28

<160> 14

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 3931
<212> DNA
<213> Plasmodium falciparum

<400> 1
taaatcttgg taataaaata aaagtgagga gtaaaactat gaacagtaaa ataaaagttg 60
tggtgaggaa gagaccactg agcgaattag aaaagaagaa aaaagatagt gatataatta 120
cagtaaaaaa caattgtacg ctttatatag atgaaccaag atataaagtg gatatgacaa 180
aatatataga aaggcatgaa tttattgtag ataaagtttt tgatgatacg gttgataatt 240
tcacagtata tgagaatacc ataaaaccat taataataga tttatatgag aatggttgtg 300
tatgttcttg ttttgcttat gggcaaacag gtagcgggaa gacttatacg atgttaggtt 360
cacaaccgta tggacagagt gataccctg gtatatttca atacgcagca ggggatatat 420
ttacctttt aaatatttat gataaagata atacgaaagg gatatttata tcattttatg 480
aaatttattg tggtaaatta tatgatttat tacaaaaacg taagatggta gcagcattag 540
aaaatgggaa aaaagaagtt gtagtaaaag atttaaaaat attaagagta ttaacaaaag 600
aagaattaat attaaaaatg atagatggtg ttttattaag aaaaattggt gttaattcac 660
aaaacgatga atcatctaga tcacatgcta tattaaatat tgatttaaaa gatataaata 720
aaaatacatc tcttggaaaa attgctttca ttgatttagc aggaagtgaa agaggagctg 780
ataccgtttc acaaaataaa caaacacaaa ccgatggagc taatattaat agatctttac 840
tagccttaaa ggaatgtatt cgagctatgg attcagataa aaatcatata cctttcagag 900
attcagaatt aactaaagtt ttaagagata tatttgtagg gaaatctaaa agtattatga 960
tagctaatat ttctcctaca attagttgtt gtgagcaaac attgaataca ttaagatatt 1020
cttcaagagt taagaacttt aaaaataaat ctacatgtat aaatgaagaa gatgatacaa 1080
ataccgaaag aattagtata ttagattcaa agggaagtga aatgaatgca tctagtatag 1140
aaaatgtagt tatcaaatca aatcatcttc tttcaaataa taataataat aaaatcaatc 1200
gtggcaaaat taatgataag atagaaagaa ataatatctt aaaaaataaa tcttttgata 1260
aacctaggga aggattcaca tcaacctttg gaaaatacag ttcacttaat gatatagaca 1320
aaataaagaa aaataagaaa aagggtttaa ttaattataa aagtacatta tataatgaca 1380
acaccataaa taaaaagcac aacaataata ataataataa taatgataat aatgataata 1440
ataatgataa taataataat aataataatg atagtagtag tatggtcaat aacatgatca 1500
atcatatgat caataataat attaacaaca atatcaatgt aaataacaac aataataata 1560
ataataataa caataatagt cataataatc atctaccaca acccaactat gcatttaccg 1620
atacatccga tttttcttcg cttgacgata tgaactgtca tttaaacaat aacgacaaaa 1680

gtatttttct acacaaaaaa aatttaagag acaacattaa attgaaaaac agaagtagtt 1740
gtgataacat tatgaacaag aagaaaaaca atcttcattt agctagacat agtgtcggaa 1800
gtaaattaac aatgttttca tatgacccac aaaaaaataa ggataataca ttttttaaaa 1860
gtaatataaa caaaatggaa gataatacac ctaaggatat actatatgag tctaggaatg 1920
tatcgaatat gaatggtaat gttttattag ggttaaacaa gaatactcat cacgatatct 1980
cgacaaagga tgaaaatcat aatgataata aaataaacaa tggtgttatt aacattataa 2040
ataatagtaa tgtgaatagt ataaataata gtaatatgaa tagtataaat aatagtaata 2100
tgaatagcaa tagtatttat aaaagtaact ataatagtaa tcagtccata tcagatgttc 2160
aaattagata cgtgaacgaa atggacacta gcaataaaaa taatgacaac atatttttg 2220
atgctatatc atgtgataat aacatgtatc ctaatataac aaataataat aataataata 2280
ataataataa taataataat aataacaata ttgatgtgga gaattataat aatcgtgatg 2340
gtacaaataa ctctatgaag ttgtatgctt ataatagtca taacttattt caacctgata 2400
ataataaaaa tacatcaaac attcagaata taaatacaaa taaaaataat caagatggga 2460
atgtaaatta ttcaatgaat ttttgccatt ataatttaaa tgataaaaat tatttgatag 2520
atttaaataa taaggaacaa aaggataaaa acatacatgg atgtgacaat aatattattc 2580
aaaataggaa cgattttgaa aaaaagaaaa aaaccaattt ttacaataat aataatattg 2640
ttatagtaaa taataatatg ggaaataaca atagtccccg tatgaaatat ggtttatgtg 2700
gtagtcatac aagtattgat aatatgaaaa ataatgaaat gaaaaataat gaaatgaaag 2760
ataatgaaat gaaagataat catataaaaa gtaacaataa taatagtagt agtagtagta 2820
gtagtaataa taatatttat aataatatta atgatgatga tacatttcaa aatgattatt 2880
gtcacaatga taataccttt actattagac gaaaaaacaa tactaatata aatagtaaca 2940
tataccaaaa tgatgatata atttatacaa taaatagttt aaatgattat atgagcaata 3000
ccctgttaca ttttaaagag aaatatacat atccaacact aagcacaaat gaagatatat 3060
ataataaaga aatggaagga aaacatataa ggctggatga tcaagataaa tatgatgata 3120
atgataataa taatgttgat aataataata aaaataatgt tgataataat gttgataata 3180
ataatgttga taataatgtt gataataatg ataaaaataa tgttgataat aataatgttg 3240
ataatgatga tgatgatgta gattttcata atataaaaaa ttttaataat aatgaatatc 3300
ttagctactt tcaaaaaaat gtagatacaa taataaataa ttgcttaaat tctttagata 3360
tttctagtat gtatgatgat acaaaagaga tactaaataa tatcctatta tctaaatata 3420
aagctgaaaa agataatgtt ataaaaaaat atattaatga agatataaaa aatatgtctt 3480
tagaagaaat cgataaaaca gctcagtcca tttatgaaaa gagaaaagta ttacttacaa 3540
aattattatt attatttaaa aaaaatgtag atacacaaat aaataatgaa acaagtgatt 3600
taagaaaaga tcttgttatg tgtcacatat gtaataataa tcctgatgat caatttcatt 3660
tttatgcata tagtagacta gaaaaagata ttattaattt aattatgtta agacaaatat 3720
ggtgtgagag tgaaaactta agactcttat atcaattctt agtagtagaa tatcaaaata 3780
aatcagcaaa ttctgtttta ttaaatgtct cttcaaataa tggtgacatt atattactta 3840
ataaaaaatt ggttcaagat aacatcaaaa attctatgga ccacaacaat atacacaaaa 3900
aataaagaat attaatataa ataagtaaat a 3931

<210> 2
<211> 1288
<212> PRT
<213> Plasmodium falciparum

<400> 2
Met Asn Ser Lys Ile Lys Val Val Val Arg Lys Arg Pro Leu Ser Glu
1 5 10 15
Leu Glu Lys Lys Lys Lys Asp Ser Asp Ile Ile Thr Val Lys Asn Asn
20 25 30
Cys Thr Leu Tyr Ile Asp Glu Pro Arg Tyr Lys Val Asp Met Thr Lys
35 40 45
Tyr Ile Glu Arg His Glu Phe Ile Val Asp Lys Val Phe Asp Asp Thr
50 55 60
Val Asp Asn Phe Thr Val Tyr Glu Asn Thr Ile Lys Pro Leu Ile Ile
65 70 75 80
Asp Leu Tyr Glu Asn Gly Cys Val Cys Ser Cys Phe Ala Tyr Gly Gln
85 90 95
Thr Gly Ser Gly Lys Thr Tyr Thr Met Leu Gly Ser Gln Pro Tyr Gly
100 105 110
Gln Ser Asp Thr Pro Gly Ile Phe Gln Tyr Ala Ala Gly Asp Ile Phe
115 120 125

```
Thr Phe Leu Asn Ile Tyr Asp Lys Asp Asn Thr Lys Gly Ile Phe Ile
    130                 135                 140
Ser Phe Tyr Glu Ile Tyr Cys Gly Lys Leu Tyr Asp Leu Leu Gln Lys
145                 150                 155                 160
Arg Lys Met Val Ala Ala Leu Glu Asn Gly Lys Lys Glu Val Val Val
                165                 170                 175
Lys Asp Leu Lys Ile Leu Arg Val Leu Thr Lys Glu Glu Leu Ile Leu
            180                 185                 190
Lys Met Ile Asp Gly Val Leu Leu Arg Lys Ile Gly Val Asn Ser Gln
        195                 200                 205
Asn Asp Glu Ser Ser Arg Ser His Ala Ile Leu Asn Ile Asp Leu Lys
    210                 215                 220
Asp Ile Asn Lys Asn Thr Ser Leu Gly Lys Ile Ala Phe Ile Asp Leu
225                 230                 235                 240
Ala Gly Ser Glu Arg Gly Ala Asp Thr Val Ser Gln Asn Lys Gln Thr
                245                 250                 255
Gln Thr Asp Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Lys Glu
            260                 265                 270
Cys Ile Arg Ala Met Asp Ser Asp Lys Asn His Ile Pro Phe Arg Asp
        275                 280                 285
Ser Glu Leu Thr Lys Val Leu Arg Asp Ile Phe Val Gly Lys Ser Lys
    290                 295                 300
Ser Ile Met Ile Ala Asn Ile Ser Pro Thr Ile Ser Cys Cys Glu Gln
305                 310                 315                 320
Thr Leu Asn Thr Leu Arg Tyr Ser Ser Arg Val Lys Asn Phe Lys Asn
                325                 330                 335
Lys Ser Thr Cys Ile Asn Glu Glu Asp Asp Thr Asn Thr Glu Arg Ile
            340                 345                 350
Ser Ile Leu Asp Ser Lys Gly Ser Glu Met Asn Ala Ser Ser Ile Glu
        355                 360                 365
Asn Val Val Ile Lys Ser Asn His Leu Leu Ser Asn Asn Asn Asn Asn
    370                 375                 380
Lys Ile Asn Arg Gly Lys Ile Asn Asp Lys Ile Glu Arg Asn Asn Ile
385                 390                 395                 400
Leu Lys Asn Lys Ser Phe Asp Lys Pro Arg Glu Gly Phe Thr Ser Thr
                405                 410                 415
Phe Gly Lys Tyr Ser Ser Leu Asn Asp Ile Asp Lys Ile Lys Lys Asn
            420                 425                 430
Lys Lys Lys Gly Leu Ile Asn Tyr Lys Ser Thr Leu Tyr Asn Asp Asn
        435                 440                 445
Thr Ile Asn Lys Lys His Asn Asn Asn Asn Asn Asn Asn Asn Asp Asn
    450                 455                 460
Asn Asp Asn Asn Asn Asp Asn Asn Asn Asn Asn Asn Asn Asp Ser Ser
465                 470                 475                 480
Ser Met Val Asn Asn Met Ile Asn His Met Ile Asn Asn Asn Ile Asn
                485                 490                 495
Asn Asn Ile Asn Val Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn
            500                 505                 510
Asn Ser His Asn Asn His Leu Pro Gln Pro Asn Tyr Ala Phe Thr Asp
        515                 520                 525
Thr Ser Asp Phe Ser Ser Leu Asp Asp Met Asn Cys His Leu Asn Asn
    530                 535                 540
Asn Asp Lys Ser Ile Phe Leu His Lys Lys Asn Leu Arg Asp Asn Ile
545                 550                 555                 560
Lys Leu Lys Asn Arg Ser Ser Cys Asp Asn Ile Met Asn Lys Lys Lys
                565                 570                 575
Asn Asn Leu His Leu Ala Arg His Ser Val Gly Ser Lys Leu Thr Met
            580                 585                 590
Phe Ser Tyr Asp Pro Gln Lys Asn Lys Asp Asn Thr Phe Phe Lys Ser
        595                 600                 605
```

```
Asn Ile Asn Lys Met Glu Asp Asn Thr Pro Lys Asp Ile Leu Tyr Glu
    610                 615                 620
Ser Arg Asn Val Ser Asn Met Asn Gly Asn Val Leu Leu Gly Leu Asn
625                 630                 635                 640
Lys Asn Thr His His Asp Ile Ser Thr Lys Asp Glu Asn His Asn Asp
                645                 650                 655
Asn Lys Ile Asn Asn Gly Val Ile Asn Ile Ile Asn Asn Ser Asn Val
            660                 665                 670
Asn Ser Ile Asn Asn Ser Asn Met Asn Ser Ile Asn Asn Ser Asn Met
        675                 680                 685
Asn Ser Asn Ser Ile Tyr Lys Ser Asn Tyr Asn Ser Asn Gln Ser Ile
    690                 695                 700
Ser Asp Val Gln Ile Arg Tyr Val Asn Glu Met Asp Thr Ser Asn Lys
705                 710                 715                 720
Asn Asn Asp Asn Ile Phe Phe Asp Ala Ile Ser Cys Asp Asn Asn Met
                725                 730                 735
Tyr Pro Asn Ile Thr Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn
            740                 745                 750
Asn Asn Asn Asn Asn Ile Asp Val Glu Asn Tyr Asn Asn Arg Asp Gly
        755                 760                 765
Thr Asn Asn Ser Met Lys Leu Tyr Ala Tyr Asn Ser His Asn Leu Phe
    770                 775                 780
Gln Pro Asp Asn Asn Lys Asn Thr Ser Asn Ile Gln Asn Ile Asn Thr
785                 790                 795                 800
Asn Lys Asn Asn Gln Asp Gly Asn Val Asn Tyr Ser Met Asn Phe Cys
                805                 810                 815
His Tyr Asn Leu Asn Asp Lys Asn Tyr Leu Ile Asp Leu Asn Asn Lys
            820                 825                 830
Glu Gln Lys Asp Lys Asn Ile His Gly Cys Asp Asn Asn Ile Ile Gln
        835                 840                 845
Asn Arg Asn Asp Phe Glu Lys Lys Lys Lys Thr Asn Phe Tyr Asn Asn
    850                 855                 860
Asn Asn Ile Val Ile Val Asn Asn Asn Met Gly Asn Asn Asn Ser Pro
865                 870                 875                 880
Arg Met Lys Tyr Gly Leu Cys Gly Ser His Thr Ser Ile Asp Asn Met
                885                 890                 895
Lys Asn Asn Glu Met Lys Asn Asn Glu Met Lys Asp Asn Glu Met Lys
            900                 905                 910
Asp Asn His Ile Lys Ser Asn Asn Asn Asn Ser Ser Ser Ser Ser Ser
        915                 920                 925
Ser Asn Asn Asn Ile Tyr Asn Asn Ile Asn Asp Asp Asp Thr Phe Gln
    930                 935                 940
Asn Asp Tyr Cys His Asn Asp Asn Thr Phe Thr Ile Arg Arg Lys Asn
945                 950                 955                 960
Asn Thr Asn Ile Asn Ser Asn Ile Tyr Gln Asn Asp Asp Ile Ile Tyr
                965                 970                 975
Thr Ile Asn Ser Leu Asn Asp Tyr Met Ser Asn Thr Leu Leu His Phe
            980                 985                 990
Lys Glu Lys Tyr Thr Tyr Pro Thr Leu Ser Thr Asn Glu Asp Ile Tyr
        995                 1000                1005
Asn Lys Glu Met Glu Gly Lys His Ile Arg Leu Asp Asp Gln Asp Lys
    1010                1015                1020
Tyr Asp Asp Asn Asp Asn Asn Asn Val Asp Asn Asn Asn Lys Asn Asn
1025                1030                1035                1040
Val Asp Asn Asn Val Asp Asn Asn Asn Val Asp Asn Asn Val Asp Asn
                1045                1050                1055
Asn Asp Lys Asn Asn Val Asp Asn Asn Asn Val Asp Asn Asp Asp Asp
            1060                1065                1070
Asp Val Asp Phe His Asn Ile Lys Asn Phe Asn Asn Asn Glu Tyr Leu
        1075                1080                1085
```

```
Ser Tyr Phe Gln Lys Asn Val Asp Thr Ile Ile Asn Asn Cys Leu Asn
    1090                1095                1100
Ser Leu Asp Ile Ser Ser Met Tyr Asp Asp Thr Lys Glu Ile Leu Asn
1105                1110                1115                1120
Asn Ile Leu Leu Ser Lys Tyr Lys Ala Glu Lys Asp Asn Val Ile Lys
                1125                1130                1135
Lys Tyr Ile Asn Glu Asp Ile Lys Asn Met Ser Leu Glu Glu Ile Asp
            1140                1145                1150
Lys Thr Ala Gln Ser Ile Tyr Glu Lys Arg Lys Val Leu Leu Thr Lys
        1155                1160                1165
Leu Leu Leu Leu Phe Lys Lys Asn Val Asp Thr Gln Ile Asn Asn Glu
    1170                1175                1180
Thr Ser Asp Leu Arg Lys Asp Leu Val Met Cys His Ile Cys Asn Asn
1185                1190                1195                1200
Asn Pro Asp Asp Gln Phe His Phe Tyr Ala Tyr Ser Arg Leu Glu Lys
                1205                1210                1215
Asp Ile Ile Asn Leu Ile Met Leu Arg Gln Ile Trp Cys Glu Ser Glu
            1220                1225                1230
Asn Leu Arg Leu Leu Tyr Gln Phe Leu Val Val Glu Tyr Gln Asn Lys
        1235                1240                1245
Ser Ala Asn Ser Val Leu Leu Asn Val Ser Ser Asn Asn Gly Asp Ile
    1250                1255                1260
Ile Leu Leu Asn Lys Lys Leu Val Gln Asp Asn Ile Lys Asn Ser Met
1265                1270                1275                1280
Asp His Asn Asn Ile His Lys Lys
                1285
```

<210> 3
<211> 990
<212> DNA
<213> Plasmodium falciparum

<400> 3
aaagttgtgg tgaggaagag accactgagc gaattagaaa agaagaaaaa agatagtgat 60
ataattacag taaaaaacaa ttgtacgctt tatatagatg aaccaagata taaagtggat 120
atgacaaaat atatagaaag gcatgaattt attgtagata aagtttttga tgatacggtt 180
gataatttca cagtatatga gaataccata aaaccattaa taatagattt atatgagaat 240
ggttgtgtat gttcttgttt tgcttatggg caaacaggta gcgggaagac ttatacgatg 300
ttaggttcac aaccgtatgg acagagtgat acccctggta tatttcaata cgcagcaggg 360
gatatattta ccttttaaa tatttatgat aaagataata cgaaagggat atttatatca 420
ttttatgaaa tttattgtgg taaattatat gatttattac aaaaacgtaa gatggtagca 480
gcattagaaa atgggaaaaa agaagttgta gtaaaagatt taaaaatatt aagagtatta 540
acaaaagaag aattaatatt aaaaatgata gatggtgttt tattaagaaa aattggtgtt 600
aattcacaaa acgatgaatc atctagatca catgctatat taaatattga tttaaaagat 660
ataaataaaa atacatctct tggaaaaatt gctttcattg atttagcagg aagtgaaaga 720
ggagctgata ccgtttcaca aaataaacaa acacaaaccg atggagctaa tattaataga 780
tctttactag ccttaaagga atgtattcga gctatggatt cagataaaaa tcatatacct 840
ttcagagatt cagaattaac taaagtttta agagatatat ttgtagggaa atctaaaagt 900
attatgatag ctaatatttc tcctacaatt agttgttgtg agcaaacatt gaatacatta 960
agatattctt caagagttaa gaactttaaa 990

<210> 4
<211> 332
<212> PRT
<213> Plasmodium falciparum

<400> 4
```
Lys Ile Lys Val Val Val Arg Lys Arg Pro Leu Ser Glu Leu Glu Lys
 1               5                  10                  15
```

```
Lys Lys Lys Asp Ser Asp Ile Ile Thr Val Lys Asn Asn Cys Thr Leu
            20                  25                  30
Tyr Ile Asp Glu Pro Arg Tyr Lys Val Asp Met Thr Lys Tyr Ile Glu
        35                  40                  45
Arg His Glu Phe Ile Val Asp Lys Val Phe Asp Asp Thr Val Asp Asn
    50                  55                  60
Phe Thr Val Tyr Glu Asn Thr Ile Lys Pro Leu Ile Ile Asp Leu Tyr
65                  70                  75                  80
Glu Asn Gly Cys Val Cys Ser Cys Phe Ala Tyr Gly Gln Thr Gly Ser
                85                  90                  95
Gly Lys Thr Tyr Thr Met Leu Gly Ser Gln Pro Tyr Gly Gln Ser Asp
            100                 105                 110
Thr Pro Gly Ile Phe Gln Tyr Ala Ala Gly Asp Ile Phe Thr Phe Leu
        115                 120                 125
Asn Ile Tyr Asp Lys Asp Asn Thr Lys Gly Ile Phe Ile Ser Phe Tyr
    130                 135                 140
Glu Ile Tyr Cys Gly Lys Leu Tyr Asp Leu Leu Gln Lys Arg Lys Met
145                 150                 155                 160
Val Ala Ala Leu Glu Asn Gly Lys Lys Glu Val Val Val Lys Asp Leu
                165                 170                 175
Lys Ile Leu Arg Val Leu Thr Lys Glu Glu Leu Ile Leu Lys Met Ile
            180                 185                 190
Asp Gly Val Leu Leu Arg Lys Ile Gly Val Asn Ser Gln Asn Asp Glu
        195                 200                 205
Ser Ser Arg Ser His Ala Ile Leu Asn Ile Asp Leu Lys Asp Ile Asn
    210                 215                 220
Lys Asn Thr Ser Leu Gly Lys Ile Ala Phe Ile Asp Leu Ala Gly Ser
225                 230                 235                 240
Glu Arg Gly Ala Asp Thr Val Ser Gln Asn Lys Gln Thr Gln Thr Asp
                245                 250                 255
Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Lys Glu Cys Ile Arg
            260                 265                 270
Ala Met Asp Ser Asp Lys Asn His Ile Pro Phe Arg Asp Ser Glu Leu
        275                 280                 285
Thr Lys Val Leu Arg Asp Ile Phe Val Gly Lys Ser Lys Ser Ile Met
    290                 295                 300
Ile Ala Asn Ile Ser Pro Thr Ile Ser Cys Cys Glu Gln Thr Leu Asn
305                 310                 315                 320
Thr Leu Arg Tyr Ser Ser Arg Val Lys Asn Phe Lys
                325                 330
```

<210> 5
<211> 1074
<212> DNA
<213> Plasmodium falciparum

<400> 5
atgaaaataa aagttgtggt gaggaagaga ccactgagcg aattagaaaa gaagaaaaaa 60
gatagtgata taattacagt aaaaaacaat tgtacgcttt atatagatga accaagatat 120
aaagtggata tgacaaaata tatagaaagg catgaattta ttgtagataa agtttttgat 180
gatacggttg ataatttcac agtatatgag aataccataa aaccattaat aatagattta 240
tatgagaatg gttgtgtatg ttcttgtttt gcttatgggc aaacaggtag cgggaagact 300
tatacgatgt taggttcaca accgtatgga cagagtgata cccctggtat atttcaatac 360
gcagcagggg atatatttac ctttttaaat atttatgata aagataatac gaaagggata 420
tttatatcat tttatgaaat ttattgtggt aaattatatg atttattaca aaaacgtaag 480
atggtagcag cattagaaaa tgggaaaaaa gaagttgtag taaaagattt aaaaatatta 540
agagtattaa caaaagaaga attaatatta aaaatgatag atggtgtttt attaagaaaa 600
attggtgtta attcacaaaa cgatgaatca tctagatcac atgctatatt aaatattgat 660
ttaaaagata taaataaaaa tacatctctt ggaaaaattg ctttcattga tttagcagga 720
agtgaaagag gagctgatac cgtttcacaa aataaacaaa cacaaaccga tggagctaat 780

attaatagat ctttactagc cttaaaggaa tgtattcgag ctatggattc agataaaaat 840
catatacctt tcagagattc agaattaact aaagttttaa gagatatatt tgtagggaaa 900
tctaaaagta ttatgatagc taatatttct cctacaatta gttgttgtga gcaaacattg 960
aatacattaa gatattcttc aagagttaag aactttaaaa ataaatctac atgtataaat 1020
gaagaagatg atacaaatac cgaaagaatt agtatattag attcaaaggg atga 1074

<210> 6
<211> 355
<212> PRT
<213> Plasmodium falciparum

<400> 6
Met Lys Ile Lys Val Val Val Arg Lys Arg Pro Leu Ser Glu Leu Glu
1 5 10 15
Lys Lys Lys Lys Asp Ser Asp Ile Ile Thr Val Lys Asn Asn Cys Thr
20 25 30
Leu Tyr Ile Asp Glu Pro Arg Tyr Lys Val Asp Met Thr Lys Tyr Ile
35 40 45
Glu Arg His Glu Phe Ile Val Asp Lys Val Phe Asp Asp Thr Val Asp
50 55 60
Asn Phe Thr Val Tyr Glu Asn Thr Ile Lys Pro Leu Ile Ile Asp Leu
65 70 75 80
Tyr Glu Asn Gly Cys Val Cys Ser Cys Phe Ala Tyr Gly Gln Thr Gly
85 90 95
Ser Gly Lys Thr Tyr Thr Met Leu Gly Ser Gln Pro Tyr Gly Gln Ser
100 105 110
Asp Thr Pro Gly Ile Phe Gln Tyr Ala Ala Gly Asp Ile Phe Thr Phe
115 120 125
Leu Asn Ile Tyr Asp Lys Asp Asn Thr Lys Gly Ile Phe Ile Ser Phe
130 135 140
Tyr Glu Ile Tyr Cys Gly Lys Leu Tyr Asp Leu Leu Gln Lys Arg Lys
145 150 155 160
Met Val Ala Ala Leu Glu Asn Gly Lys Lys Glu Val Val Val Lys Asp
165 170 175
Leu Lys Ile Leu Arg Val Leu Thr Lys Glu Glu Leu Ile Leu Lys Met
180 185 190
Ile Asp Gly Val Leu Leu Arg Lys Ile Gly Val Asn Ser Gln Asn Asp
195 200 205
Glu Ser Ser Arg Ser His Ala Ile Leu Asn Ile Asp Leu Lys Asp Ile
210 215 220
Asn Lys Asn Thr Ser Leu Gly Lys Ile Ala Phe Ile Asp Leu Ala Gly
225 230 235 240
Ser Glu Arg Gly Ala Asp Thr Val Ser Gln Asn Lys Gln Thr Gln Thr
245 250 255
Asp Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Lys Glu Cys Ile
260 265 270
Arg Ala Met Asp Ser Asp Lys Asn His Ile Pro Phe Arg Asp Ser Glu
275 280 285
Leu Thr Lys Val Leu Arg Asp Ile Phe Val Gly Lys Ser Lys Ser Ile
290 295 300
Met Ile Ala Asn Ile Ser Pro Thr Ile Ser Cys Cys Glu Gln Thr Leu
305 310 315 320
Asn Thr Leu Arg Tyr Ser Ser Arg Val Lys Asn Phe Lys Asn Lys Ser
325 330 335
Thr Cys Ile Asn Glu Glu Asp Asp Thr Asn Thr Glu Arg Ile Ser Ile
340 345 350
Leu Asp Ser
355

<210> 7
<211> 1086
<212> DNA
<213> Plasmodium falciparum

<400> 7
atgaaaataa aagttgtggt gaggaagaga ccactgagcg aattagaaaa gaagaaaaaa 60
gatagtgata taattacagt aaaaaacaat tgtacgcttt atatagatga accaagatat 120
aaagtggata tgacaaaata tatagaaagg catgaattta ttgtagataa agtttttgat 180
gatacggttg ataatttcac agtatatgag aataccataa aaccattaat aatagattta 240
tatgagaatg gttgtgtatg ttcttgtttt gcttatgggc aaacaggtag cgggaagact 300
tatacgatgt taggttcaca accgtatgga cagagtgata cccctggtat atttcaatac 360
gcagcagggg atatatttac ctttttaaat atttatgata aagataatac gaaagggata 420
tttatatcat tttatgaaat ttattgtggt aaattatatg atttattaca aaaacgtaag 480
atggtagcag cattagaaaa tgggaaaaaa gaagttgtag taaaagattt aaaaatatta 540
agagtattaa caaaagaaga attaatatta aaaatgatag atggtgtttt attaagaaaa 600
attggtgtta attcacaaaa cgatgaatca tctagatcac atgctatatt aaatattgat 660
ttaaaagata taaataaaaa tacatctctt ggaaaaattg ctttcattga tttagcagga 720
agtgaaagag gagctgatac cgtttcacaa aataaacaaa cacaaaccga tggagctaat 780
attaatagat ctttactagc cttaaaggaa tgtattcgag ctatggattc agataaaaat 840
catatacctt tcagagattc agaattaact aaagttttaa gagatatatt tgtagggaaa 900
tctaaaagta ttatgatagc taatatttct cctacaatta gttgttgtga gcaaacattg 960
aatacattaa gatattcttc aagagttaag aacaagggca attcgaagct tgaaggtaag 1020
cctatcccta accctctcct cggtctcgat tctacgcgta ccggtcatca tcaccatcac 1080
cattga 1086

<210> 8
<211> 361
<212> PRT
<213> Plasmodium falciparum

<400> 8
Met Lys Ile Lys Val Val Val Arg Lys Arg Pro Leu Ser Glu Leu Glu
 1   5   10   15
Lys Lys Lys Lys Asp Ser Asp Ile Ile Thr Val Lys Asn Asn Cys Thr
            20   25   30
Leu Tyr Ile Asp Glu Pro Arg Tyr Lys Val Asp Met Thr Lys Tyr Ile
        35   40   45
Glu Arg His Glu Phe Ile Val Asp Lys Val Phe Asp Asp Thr Val Asp
    50   55   60
Asn Phe Thr Val Tyr Glu Asn Thr Ile Lys Pro Leu Ile Ile Asp Leu
65   70   75   80
Tyr Glu Asn Gly Cys Val Cys Ser Cys Phe Ala Tyr Gly Gln Thr Gly
                85   90   95
Ser Gly Lys Thr Tyr Thr Met Leu Gly Ser Gln Pro Tyr Gly Gln Ser
            100   105   110
Asp Thr Pro Gly Ile Phe Gln Tyr Ala Ala Gly Asp Ile Phe Thr Phe
        115   120   125
Leu Asn Ile Tyr Asp Lys Asp Asn Thr Lys Gly Ile Phe Ile Ser Phe
    130   135   140
Tyr Glu Ile Tyr Cys Gly Lys Leu Tyr Asp Leu Leu Gln Lys Arg Lys
145   150   155   160
Met Val Ala Ala Leu Glu Asn Gly Lys Lys Glu Val Val Val Lys Asp
                165   170   175
Leu Lys Ile Leu Arg Val Leu Thr Lys Glu Glu Leu Ile Leu Lys Met
            180   185   190
Ile Asp Gly Val Leu Leu Arg Lys Ile Gly Val Asn Ser Gln Asn Asp
        195   200   205
Glu Ser Ser Arg Ser His Ala Ile Leu Asn Ile Asp Leu Lys Asp Ile
    210   215   220

```
Asn Lys Asn Thr Ser Leu Gly Lys Ile Ala Phe Ile Asp Leu Ala Gly
225                 230                 235                 240
Ser Glu Arg Gly Ala Asp Thr Val Ser Gln Asn Lys Gln Thr Gln Thr
                245                 250                 255
Asp Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Lys Glu Cys Ile
            260                 265                 270
Arg Ala Met Asp Ser Asp Lys Asn His Ile Pro Phe Arg Asp Ser Glu
        275                 280                 285
Leu Thr Lys Val Leu Arg Asp Ile Phe Val Gly Lys Ser Lys Ser Ile
    290                 295                 300
Met Ile Ala Asn Ile Ser Pro Thr Ile Ser Cys Cys Glu Gln Thr Leu
305                 310                 315                 320
Asn Thr Leu Arg Tyr Ser Ser Arg Val Lys Asn Lys Gly Asn Ser Lys
                325                 330                 335
Leu Glu Gly Lys Pro Ile Pro Asn Pro Leu Leu Gly Leu Asp Ser Thr
            340                 345                 350
Arg Thr Gly His His His His His His
        355                 360
```

<210> 9
<211> 987
<212> DNA
<213> Plasmodium falciparum

<400> 9
ataaaagttg tggtgaggaa gagaccactg agcgaattag aaaagaagaa aaaagatagt 60
gatataatta cagtaaaaaa caattgtacg ctttatatag atgaaccaag atataaagtg 120
gatatgacaa aatatataga aaggcatgaa tttattgtag ataaagtttt tgatgatacg 180
gttgataatt tcacagtata tgagaatacc ataaaaccat taataataga tttatatgag 240
aatggttgtg tatgttcttg ttttgcttat gggcaaacag gtagcgggaa gacttatacg 300
atgttaggtt cacaaccgta tggacagagt gataccctg gtatatttca atacgcagca 360
ggggatatat ttacctttt aaatatttat gataaagata atacgaaagg gatatttata 420
tcattttatg aaatttattg tggtaaatta tatgatttat tacaaaaacg taagatggta 480
gcagcattag aaaatgggaa aaaagaagtt gtagtaaaag atttaaaaat attaagagta 540
ttaacaaaag aagaattaat attaaaaatg atagatggtg ttttattaag aaaaattggt 600
gttaattcac aaaacgatga atcatctaga tcacatgcta tattaaatat tgatttaaaa 660
gatataaata aaaatacatc tcttggaaaa attgctttca ttgatttagc aggaagtgaa 720
agaggagctg ataccgtttc acaaaataaa caaacacaaa ccgatggagc taatattaat 780
agatctttac tagccttaaa ggaatgtatt cgagctatgg attcagataa aaatcatata 840
cctttcagag attcagaatt aactaaagtt ttaagagata tatttgtagg gaaatctaaa 900
agtattatga tagctaatat ttctcctaca attagttgtt gtgagcaaac attgaataca 960
ttaagatatt cttcaagagt taagaac 987

<210> 10
<211> 332
<212> PRT
<213> Plasmodium falciparum

<400> 10
```
Ile Lys Val Val Val Arg Lys Arg Pro Leu Ser Glu Leu Glu Lys Lys
 1               5                  10                  15
Lys Lys Asp Ser Asp Ile Ile Thr Val Lys Asn Asn Cys Thr Leu Tyr
            20                  25                  30
Ile Asp Glu Pro Arg Tyr Lys Val Asp Met Thr Lys Tyr Ile Glu Arg
        35                  40                  45
His Glu Phe Ile Val Asp Lys Val Phe Asp Asp Thr Val Asp Asn Phe
    50                  55                  60
Thr Val Tyr Glu Asn Thr Ile Lys Pro Leu Ile Ile Asp Leu Tyr Glu
65                  70                  75                  80
Asn Gly Cys Val Cys Ser Cys Phe Ala Tyr Gly Gln Thr Gly Ser Gly
```

```
                85                  90                  95
Lys Thr Tyr Thr Met Leu Gly Ser Gln Pro Tyr Gly Gln Ser Asp Thr
            100                 105                 110
Pro Gly Ile Phe Gln Tyr Ala Ala Gly Asp Ile Phe Thr Phe Leu Asn
        115                 120                 125
Ile Tyr Asp Lys Asp Asn Thr Lys Gly Ile Phe Ile Ser Phe Tyr Glu
    130                 135                 140
Ile Tyr Cys Gly Lys Leu Tyr Asp Leu Leu Gln Lys Arg Lys Met Val
145                 150                 155                 160
Ala Ala Leu Glu Asn Gly Lys Lys Glu Val Val Val Lys Asp Leu Lys
                165                 170                 175
Ile Leu Arg Val Leu Thr Lys Glu Glu Leu Ile Leu Lys Met Ile Asp
            180                 185                 190
Gly Val Leu Leu Arg Lys Ile Gly Val Asn Ser Gln Asn Asp Glu Ser
        195                 200                 205
Ser Arg Ser His Ala Ile Leu Asn Ile Asp Leu Lys Asp Ile Asn Lys
    210                 215                 220
Asn Thr Ser Leu Gly Lys Ile Ala Phe Ile Asp Leu Ala Gly Ser Glu
225                 230                 235                 240
Arg Gly Ala Asp Thr Val Ser Gln Asn Lys Gln Thr Gln Thr Asp Gly
                245                 250                 255
Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Lys Glu Cys Ile Arg Ala
            260                 265                 270
Met Asp Ser Asp Lys Asn His Ile Pro Phe Arg Asp Ser Glu Leu Thr
        275                 280                 285
Lys Val Leu Arg Asp Ile Phe Val Gly Lys Ser Lys Ser Ile Met Ile
    290                 295                 300
Ala Asn Ile Ser Pro Thr Ile Ser Cys Cys Glu Gln Thr Leu Asn Thr
305                 310                 315                 320
Leu Arg Tyr Ser Ser Arg Val Lys Asn Phe Lys Asn
                325                 330
```

<210> 11
<211> 981
<212> DNA
<213> Plasmodium falciparum

<400> 11
gttgtggtga ggaagagacc actgagcgaa ttagaaaaga agaaaaaaga tagtgatata 60
attacagtaa aaaacaattg tacgctttat atagatgaac caagatataa agtggatatg 120
acaaaatata tagaaaggca tgaatttatt gtagataaag tttttgatga tacggttgat 180
aatttcacag tatatgagaa taccataaaa ccattaataa tagatttata tgagaatggt 240
tgtgtatgtt cttgttttgc ttatgggcaa acaggtagcg ggaagactta tacgatgtta 300
ggttcacaac cgtatggaca gagtgatacc cctggtatat ttcaatacgc agcaggggat 360
atatttacct ttttaaatat ttatgataaa gataatacga aagggatatt tatatcattt 420
tatgaaattt attgtggtaa attatatgat ttattacaaa aacgtaagat ggtagcagca 480
ttagaaaatg ggaaaaaaga agttgtagta aaagatttaa aaatattaag agtattaaca 540
aaagaagaat taatattaaa aatgatagat ggtgttttat taagaaaaat tggtgttaat 600
tcacaaaacg atgaatcatc tagatcacat gctatattaa atattgattt aaaagatata 660
aataaaaata catctcttgg aaaaattgct ttcattgatt tagcaggaag tgaaagagga 720
gctgataccg tttcacaaaa taaacaaaca caaaccgatg gagctaatat taatagatct 780
ttactagcct taaaggaatg tattcgagct atggattcag ataaaaatca tataccttttc 840
agagattcag aattaactaa agttttaaga gatatatttg tagggaaatc taaaagtatt 900
atgatagcta atatttctcc tacaattagt tgttgtgagc aaacattgaa tacattaaga 960
tattcttcaa gagttaagaa c 981

<210> 12
<211> 329
<212> PRT
<213> Plasmodium falciparum

<400> 12
Ile Lys Val Val Val Arg Lys Arg Pro Leu Ser Glu Leu Glu Lys Lys
1 5 10 15
Lys Lys Asp Ser Asp Ile Ile Thr Val Lys Asn Asn Cys Thr Leu Tyr
20 25 30
Ile Asp Glu Pro Arg Tyr Lys Val Asp Met Thr Lys Tyr Ile Glu Arg
35 40 45
His Glu Phe Ile Val Asp Lys Val Phe Asp Asp Thr Val Asp Asn Phe
50 55 60
Thr Val Tyr Glu Asn Thr Ile Lys Pro Leu Ile Ile Asp Leu Tyr Glu
65 70 75 80
Asn Gly Cys Val Cys Ser Cys Phe Ala Tyr Gly Gln Thr Gly Ser Gly
85 90 95
Lys Thr Tyr Thr Met Leu Gly Ser Gln Pro Tyr Gly Gln Ser Asp Thr
100 105 110
Pro Gly Ile Phe Gln Tyr Ala Ala Gly Asp Ile Phe Thr Phe Leu Asn
115 120 125
Ile Tyr Asp Lys Asp Asn Thr Lys Gly Ile Phe Ile Ser Phe Tyr Glu
130 135 140
Ile Tyr Cys Gly Lys Leu Tyr Asp Leu Leu Gln Lys Arg Lys Met Val
145 150 155 160
Ala Ala Leu Glu Asn Gly Lys Lys Glu Val Val Val Lys Asp Leu Lys
165 170 175
Ile Leu Arg Val Leu Thr Lys Glu Glu Leu Ile Leu Lys Met Ile Asp
180 185 190
Gly Val Leu Leu Arg Lys Ile Gly Val Asn Ser Gln Asn Asp Glu Ser
195 200 205
Ser Arg Ser His Ala Ile Leu Asn Ile Asp Leu Lys Asp Ile Asn Lys
210 215 220
Asn Thr Ser Leu Gly Lys Ile Ala Phe Ile Asp Leu Ala Gly Ser Glu
225 230 235 240
Arg Gly Ala Asp Thr Val Ser Gln Asn Lys Gln Thr Gln Thr Asp Gly
245 250 255
Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Lys Glu Cys Ile Arg Ala
260 265 270
Met Asp Ser Asp Lys Asn His Ile Pro Phe Arg Asp Ser Glu Leu Thr
275 280 285
Lys Val Leu Arg Asp Ile Phe Val Gly Lys Ser Lys Ser Ile Met Ile
290 295 300
Ala Asn Ile Ser Pro Thr Ile Ser Cys Cys Glu Gln Thr Leu Asn Thr
305 310 315 320
Leu Arg Tyr Ser Ser Arg Val Lys Asn
325

<210> 13
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:sense primer

<400> 13
atgataaaag ttgtggtgag gaag 24

<210> 14
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:antisense primer

<400> 14
gttcttaact cttgaagaat atcttaatg 29